



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 198439

TO: Ginny Portner
Location: REM-3C18
Art Unit: 1645
Friday, August 11, 2006

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: 571-272-2527

Case Serial Number: 09/904994

Paul.schulwitz@uspto.gov

Search Notes

Examiner Portner,

Please review the attached search results.

If you have any questions or if you would like to refine the search query, please feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
REM-1A65
571-272-2527

Schulwitz, Paul

From: Portner, Ginny
Sent: Friday, August 04, 2006 11:32 AM
To: Schulwitz, Paul
Subject: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached
Importance: High

Thanks for your help with this search issue. I am sending the claims to you now (Snagit training WORKed!!!!!!!)
Once I check the Spec for a program name etc I will send you another email. ginny

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

Claims:
23
34
40

72798

Schulwitz, Paul

From: Portner, Ginny
Sent: Friday, August 04, 2006 11:44 AM
To: Schulwitz, Paul
Subject: FW: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached
Importance: High

Paul, From the Specification, I found that they sued a computer program called "Align Plus for Windows, available from Scientific and Educational Software, P.O. Box 72045, Durham, NC 27722-2045 USA. I hope that this information helps. ginny

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

-----Original Message-----

From: Portner, Ginny
Sent: Friday, August 04, 2006 11:32 AM
To:Subject: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached
Importance: High

Thanks for your help with this search issue. I am sending the claims to you now (Snagit training WORKed!!!!!!!)
Once I check the Spec for a program name etc I will send you another email. ginny

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02; Mail Box 3C18
(571) 272-0862

8/7/06

Schulwitz, Paul

From: Portner, Ginny
Sent: Friday, August 04, 2006 12:05 PM
To: Schulwitz, Paul
Subject: FW: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached
Importance: High

Paul, After looking at all of the claims, I found that they claims SEQ ID Nos 1, 2 and 3. with the nucleic acid parameters of claim 23 previously emailed, and the polypeptide sequences with a different set of parameters as set forth in emailed claim 34. Thanks!!!

Ginny Portner

Remsen Building

Art Unit 1645

Room E03, B02; Mail Box 3C18

(571) 272-0862

-----Original Message-----

From: Portner, Ginny

Sent: Friday, August 04, 2006 11:44 AM

To: Subject: FW: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached

Importance: High

Paul, From the Specification, I found that they sued a computer program called "Align Plus for Windows, available from Scientific and Educational Software, P.O. Box 72045, Durham, NC 27722-2045 USA. I hope that this information helps. ginny

Ginny Portner

Remsen Building

Art Unit 1645

Room E03, B02; Mail Box 3C18

(571) 272-0862

-----Original Message-----

From: Portner, Ginny

Sent: Friday, August 04, 2006 11:32 AM

To:Subject: 09/904994 nucleic and amino acid sequence search with specific gap values; claims attached

Importance: High

Thanks for your help with this search issue. I am sending the claims to you now (Snagit training WORKed!!!!!!!)
Once I check the Spec for a program name etc I will send you another email. ginny

Ginny Portner

Remsen Building

Art Unit 1645

Room E03, B02; Mail Box 3C18

(571) 272-0862

8/7/06

In the claims:

Claims 1 - 22 (Canceled)

23. (Currently amended) A nucleic acid molecule comprising a nucleotide sequence encoding two urease X and urease Y subunit polypeptides of a urease complex such as expressed by *Helicobacter felis*, said nucleotide sequence having at least 85% homology with SEQ ID NO: 1 or a part thereof encoding at least an immunogenic fragment of one of said subunits, said part having a length of at least 40 ~~70~~ nucleotides.

34. (Currently amended) An isolated *Helicobacter felis* urease X subunit polypeptide, said polypeptide comprising an amino acid sequence that is at least 85% homologous to SEQ ID NO: 2, or an immunogenic fragment of said polypeptide having a length of at least 40 ~~70~~ amino acids, wherein said immunogenic fragment induces an immune response against an urease X subunit in ureaseXY.

Amendments to the Claims:

min length = 70

1. – 22. (cancelled)

23. (currently amended) **[[A]] An isolated** nucleic acid molecule comprising

(i) a nucleotide sequence encoding urease X and urease Y subunit polypeptides of a urease complex expressed by *Helicobacter felis*, said nucleotide sequence having at least 85% homology with SEQ ID NO: 1 as determined over a global alignment wherein the mismatch value is 1, the open gap value is 3 and the extended gap value is 3; or

(ii) a part ~~thereof~~ of (i) encoding at least an immunogenic fragment of one of said subunits, said part having a length of at least 70 nucleotides.

24. – 25. (cancelled)

26. (original) The nucleic acid molecule of claim 23, which encodes one or both of the urease X subunit polypeptide and urease Y subunit polypeptide.

27. (cancelled)

28. (currently amended) The nucleic acid molecule of claim 23, wherein the nucleotide sequence has at least 94% homology with SEQ ID No: 1 as determined over a global alignment wherein the mismatch value is 1, the open gap value is 3 and the extended gap value is 3.

29. (cancelled)

30. (previously presented) An isolated and purified DNA fragment comprising a nucleotide sequence according to claim 23.

31. (original) A recombinant DNA molecule comprising a nucleotide sequence according to Claim 23 under the control of a functionally linked promoter.

32. (original) A live recombinant carrier comprising a recombinant DNA molecule of claim 31.

33. (original) A host cell comprising a nucleic acid molecule of claim 23, a DNA fragment of claim 30, a recombinant DNA molecule of claim 31 or a live recombinant carrier of claim 32.

34. (currently amended) An isolated *Helicobacter felis* urease X subunit polypeptide, said polypeptide comprising

(a) an amino acid sequence that is at least 85% homologous to SEQ ID NO: 2 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1: [[,]] or

(b) an immunogenic fragment of (a) ~~said polypeptide~~ having a length of at least 70 amino acids, wherein said immunogenic fragment induces an immune response against [[an]] a urease X subunit in ureaseXY.

35. -36. (cancelled)

37. (currently amended) The polypeptide of claim 34, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least 90% homologous to SEQ ID NO: 2 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1: [[,]] or

(ii) an immunogenic fragment of ~~said polypeptide which~~ (i) having a length of at least 70 amino acids;
wherein said immunogenic fragment induces an immune response against [[an]] a urease X subunit in ureaseXY.

38. (currently amended) The polypeptide of claim ~~[[37]]~~ 34, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least 94% homologous to SEQ ID NO: 2 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1; [[,]] or

(ii) an immunogenic fragment of ~~said polypeptide which~~ (i) having a length of at least 70 amino acids;
wherein said immunogenic fragment induces an immune response against [[an]] a urease X subunit in ureaseXY.

39. (currently amended) The polypeptide of claim ~~[[31]]~~ 34, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least ~~[[97%]]~~ 99% homologous to SEQ ID NO: 2 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1; [[,]] or

(ii) an immunogenic fragment of ~~said polypeptide which~~ (i) having a length of at least 70 amino acids;
wherein said immunogenic fragment induces an immune response against [[an]] a urease X subunit in ureaseXY.

40. (currently amended) An isolated *Helicobacter felis* urease Y subunit polypeptide, said polypeptide comprising

(a) an amino acid sequence that is at least ~~[[85%]]~~ 86% homologous to SEQ ID NO: 3 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1; [[,]] or

(b) an immunogenic fragment of (a) ~~said polypeptide~~ having a length of at least 70 amino acids, wherein said immunogenic fragment induces an immune response against [[an]] a ureaseY subunit in ureaseXY.

41.- 43. (cancelled)

44. (currently amended) The polypeptide of claim 40, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least [[94%]] 98% homologous to SEQ ID NO: 3 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1: [[,]] or

(ii) an immunogenic fragment of said polypeptide, which (i) having a length of at least 70 amino acids;

wherein said immunogenic fragment induces an immune response against [[an]] a urease Y subunit in ureaseXY.

45. (cancelled)

46. (previously presented) An immunogenic composition comprising an immunogenically effective amount of a nucleic acid molecule of claim 23 and a pharmaceutically acceptable carrier.

47. (previously presented) The immunogenic composition of claim 46, further comprising an adjuvant.

48. (previously presented) The immunogenic composition of claim 46, further comprising an additional antigen derived from a virus or microorganism which is pathogenic to mammals.

49. (previously presented) The immunogenic composition of claim 48, wherein said virus or microorganism pathogenic to mammals is selected from the group consisting of Feline Infectious Peritonitis virus, Feline Immune deficiency virus, Canine Parvovirus, Feline Parvovirus, Distemper virus, Adenovirus, Calicivirus, *Bordetella bronchiseptica*, *Borrelia burgdorferi*,

Leptospira interrogans, *Chlamydia* and *Bartonella henseli*.

50. (previously presented) An immunogenic composition for combating *Helicobacter felis* infections, comprising antibodies against a polypeptide of claims 34 or 40.

51.- 56. (cancelled)

57. (previously presented) An immunogenic composition, comprising an immunogenically effective amount of a polypeptide according to Claim 34 and a pharmaceutically acceptable carrier.

58. (previously presented) An immunogenic composition, comprising an immunogenically effective amount of a polypeptide according to claim 40 and pharmaceutically acceptable carrier.

59. (new) The polypeptide of claim 40, wherein said polypeptide comprises:

(i) an amino acid sequence that is at least 99% homologous to SEQ ID NO: 3 as determined over a global alignment wherein the mismatch value is 2, the open gap value is 4 and the extended gap value is 1; or

(ii) an immunogenic fragment of (i) having a length of at least 70 amino acids; wherein said immunogenic fragment induces an immune response against a urease Y subunit in ureaseXY.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:13:55 ; Search time 16252 Seconds
(without alignments)
11343.860 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2880.6

Sequence: 1 rrgagatttccacatt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY_NUC

Gapop 3.0 , Gapext 3.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_env.*

2: gb_pat.*

3: gb_ph.*

4: gb_pl.*

5: gb_pr.*

6: gb_ro.*

7: gb_sts.*

8: gb_sy.*

9: gb_un.*

10: gb_vi.*

11: gb_ov.*

12: gb_htg.*

13: gb_in.*

14: gb_om.*

15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2880.6	100.0	2883	2	BD185302 Helicobac
2	2880.6	100.0	2883	2	AX356683 Sequence
3	2228	77.3	2452	2	BD185306
4	2228	77.3	2452	2	AX356695 Sequence
5	2158.6	74.9	2405	2	BD185303
6	2158.6	74.9	2405	2	AX356686 Sequence
7	2158.6	74.9	2407	2	BD185305
8	2158.6	74.9	2407	2	AX356692 Sequence
9	1948	67.6	2183	2	BD185304
10	1948	67.6	2183	2	AX356689
11	1135.6	39.4	8406	15	AF330621 Helicobac
12	1099.8	38.2	2664	15	HECUREASE
13	1040.8	36.1	110000	15	AE001439.00
14	1029.2	35.7	110000	15	AE000511.00
15	1024.4	35.6	5100	15	HECUREABCD
16	1022	35.5	7088	15	AB032429
17	1021.4	35.5	2735	2	AR054309
18	1008.8	35.0	2619	2	A41006 Sequence 1

19	1008.8	35.0	2619	2	A72509	A72509 Sequence 1
20	1008.8	35.0	2619	2	A73285	A73285 Sequence 1
21	1008.8	35.0	2619	2	AR061045	AR061045 Sequence
22	1008.8	35.0	2619	2	AR161678	AR161678 Sequence
23	1008	35.0	2746	15	HPHPUABG	X17079 Helicobacte
24	1008	35.0	2767	2	A31515	A31515 H.pylori ur
25	1005.8	34.9	4824	2	AR349305	AR349305 Sequence
26	1005.8	34.9	4824	2	AR349306	AR349306 Sequence
27	1005.4	34.9	2430	15	AY227442	AY227442 Helicobac
28	1004	34.9	2619	15	HFURE	X69080 H.felis ure
29	963.8	33.5	6126	15	AF332656	AF332656 Helicobac
30	963.8	33.5	110000	15	AE017125.03	Continuation (4 of
31	894.6	31.1	1704	15	AY714224	AY714224 Helicobac
32	885.8	30.8	1815	2	BD092530	BD092530 Identific
33	884.8	30.7	1710	15	AY295085	AY295085 Helicobac
34	881.8	30.6	1707	2	CQ903244	CQ903244 Sequence
35	881.8	30.6	1707	2	CQ903245	CQ903245 Sequence
36	880.2	30.6	1704	15	DO141576	DO141576 Helicobac
37	880	30.5	1710	15	AF352376	AF352376 Helicobac
38	879.8	30.5	1680	15	AF479026	AF479026 Helicobac
39	878.4	30.5	1710	2	A07398	A07398 C.jejuni ge
40	878.4	30.5	1710	2	A08818	A08818 C.jejuni DN
41	776.6	27.0	10118	15	BSUEROP	Y08559 Bacillus su
42	776.6	27.0	200690	15	BSUB0019	Z99122 Bacillus su
43	775.6	26.9	110000	15	BA000004.02	Continuation (3 of
44	714.6	24.8	4723	15	AB201709	AB201709 Campyloba
45	705.2	24.5	110000	15	CP000020.07	Continuation (8 of

ALIGNMENTS

RESULT 1	BD185302	2883 bp	DNA	linear	PAT 17-JUN-2003
LOCUS	BD185302	Helicobacter felis vaccine.			
DEFINITION	BD185302	Helicobacter felis vaccine.			
ACCESSION	BD185302.1	GI:31877502			
VERSION	JP 2002355054-A/1.				
KEYWORDS	JP 2002355054-A/1.				
SOURCE	Helicobacter felis				
ORGANISM	Helicobacter felis				
REFERENCE	1 (bases 1 to 2883)				
AUTHORS	Kusters, J.G. and Cattoli, G.				
TITLE	Helicobacter felis vaccine				
JOURNAL	Patent: JP 2002355054-A 1 10-DEC-2002;				
COMMENT	AKZO NOBEL NV				
	OS Helicobacter felis				
	PN JP 2002355054-A/1				
	PD 10-DEC-2002				
	PF 16-JUL-2001 JP 2001214711				
	PI 17-JUL-2000 EP 00202565.8				
	PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI				
	PC				
	C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC				
	A61K39/23				
	PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,				
	PC C12N1/19,				
	PC				
	C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC				
	53,				
	PC				
	G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1/01),(C12Q1/68, PC				
	C12R1/01),				
	PC C12N15/00,C12N5/00,A61K37/02				
	CC Helicobacter felis vaccine				
	FF Key Location/Qualifiers				
	FT CDS (206) . . (886)				
	FT CDS (897) . . (2603).				
FEATURES	Location/Qualifiers				
source	1..2883				
	/organism="Helicobacter felis"				

```
/mol_type="genomic DNA"
/db_xref="taxon:214"

ORIGIN
Query Match      100.0%; Score 2880.6; DB 2; Length 2883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGGAGATTTCCARCACTTCAAGCACATATTGATCCTGTGTGTTGGGTGGTAAATTCR 60
DB 1 RGGAGATTTTCARCACTTCAAGCACATATTGATCCTGTGTGTTGGGTGGTAAATTCR 60

QY 61 ACTTGTAAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA 120
DB 61 ACTTGTAAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA 120

QY 121 TTACTTATTAATAAAGTTAATAAAGTAACGAAATAGGACTATAATCCATTGCCTT 180
DB 121 TTACTTATTAATAAAGTTAATAAAGTAACGAAATAGGACTATAATCCATTGCCTT 180

QY 181 TAAATTTAAACAAGAGTAAATAGGTGAATCATCACCCAAAGAGCAAGAAAGTTCTT 240
DB 181 TAAATTTAAACAAGAGTAAATAGGTGAATCATCACCCAAAGAGCAAGAAAGTTCTT 240

QY 241 GTTATATTATCGCGCGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACA 300
DB 241 GTTATATTATCGCGCGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACA 300

QY 301 ACCCGAAGCCATTGCTTACATTTAGTGCCTATATTATGACCAAGCGCGCTGGAAAAA 360
DB 301 ACCCGAAGCCATTGCTTACATTTAGTGCCTATATTATGACCAAGCGCGCTGGAAAAA 360

QY 361 AACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGGAAAAAGATGAATGCC 420
DB 361 AACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGGAAAAAGATGAATGCC 420

QY 421 CGGGGTGGTAAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCCTGATGTACGAA 480
DB 421 CGGGGTGGTAAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCCTGATGTACGAA 480

QY 481 ACTTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACCTTCAAAGCGGCGAAGTCAA 540
DB 481 ACTTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACCTTCAAAGCGGCGAAGTCAA 540

QY 541 ATTTGGTTGCGATAAAGACATCGAGCTCAATGCAAGGCAAGAAAGTAAACCGAATTTGAGGT 600
DB 541 ATTTGGTTGCGATAAAGACATCGAGCTCAATGCAAGGCAAGAAAGTAAACCGAATTTGAGGT 600

QY 601 TACTAAAGAAGGCTAAATCCTTGTCATGTGGGTAGCCATTTCCACTTTTGGAAAGCTAA 660
DB 601 TACTAAAGAAGGCTAAATCCTTGTCATGTGGGTAGCCATTTCCACTTTTGGAAAGCTAA 660

QY 661 CAAGGCACTAAATTCGATCTGAAAGAGCCTATGCGAAAGCCTAGATATTCCTCTGG 720
DB 661 CAAGGCACTAAATTCGATCTGAAAGAGCCTATGCGAAAGCCTAGATATTCCTCTGG 720

QY 721 CAACACGCTACGCTATGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGG 780
DB 721 CAACACGCTACGCTATGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGG 780

QY 781 CAGTAAATAAAGTATGTCATGAACCGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
DB 781 CAGTAAATAAAGTATGTCATGAACCGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840

QY 841 ACATAAAGCGCTTCACAGGCGAAATCTCAGGATTTATCAAGTAAGAGAGACTCCCATGA 900
DB 841 ACATAAAGCGCTTCACAGGCGAAATCTCAGGATTTATCAAGTAAGAGAGACTCCCATGA 900

QY 901 AAATGAAAAACAAGAAATATGTAATACCTACGGAACCCCAAGAGCGATAAAGTGCCT 960
DB 901 AAATGAAAAACAAGAAATATGTAATACCTACGGAACCCCAAGAGCGATAAAGTGCCT 960

QY 961 TAGGAGATACCGATCTTTGGGCAGAGTAAAGATGATATACCACTTATGGCGAAGAAC 1020
```

```
DB 961 TAGGAGATACCGATCTTTGGGCAGAGTAGAAACATGACTATATACCCTATGGCGAAGAAC 1020

QY 1021 TTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATG 1080
DB 1021 TTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATG 1080

QY 1081 AAAACACCTTAGATTTAGTCATCATCAACCGCATGATTAATCGACTACACCGGATTTACA 1140
DB 1081 AAAACACCTTAGATTTAGTCATCATCAACCGCATGATTAATCGACTACACCGGATTTACA 1140

QY 1141 AAGCCGACATTTGGGATTAATAACGCGCAATCCATGGCATTCGAGGAGGAGGAAACAAG 1200
DB 1141 AAGCCGACATTTGGGATTAATAACGCGCAATCCATGGCATTCGAGGAGGAGGAAACAAG 1200

QY 1201 ACATGCAAGATGGCGCTTAAGCCCTCATATGGTCTGGGTGTGGGCACAGAAAGCACTAGCAG 1260
DB 1201 ACATGCAAGATGGCGCTTAAGCCCTCATATGGTCTGGGTGTGGGCACAGAAAGCACTAGCAG 1260

QY 1261 GGGAAAGTATGATTAATTAACCGCTGGGGGAATCGATTCACACCCCACTTCTTCTCCAC 1320
DB 1261 GGGAAAGTATGATTAATTAACCGCTGGGGGAATCGATTCACACCCCACTTCTTCTCCAC 1320

QY 1321 AACCAATCCCTACCGCTCTAGCCCAATGGCTTACCAACCATGTTTGGAGGCGGCACAGGTC 1380
DB 1321 AACCAATCCCTACCGCTCTAGCCCAATGGCTTACCAACCATGTTTGGAGGCGGCACAGGTC 1380

QY 1381 CTGTAGATGGCAAGATCGACTATCATCTCCGGGCAATGGAACCTTGCACCGCATGT 1440
DB 1381 CTGTAGATGGCAAGATCGACTATCATCTCCGGGCAATGGAACCTTGCACCGCATGT 1440

QY 1441 TGCGCGCAGCAGAAAGATTTCTTGAATGTGGGCTTTTGGGCAAGCAATAGCTCTA 1500
DB 1441 TGCGCGCAGCAGAAAGATTTCTTGAATGTGGGCTTTTGGGCAAGCAATAGCTCTA 1500

QY 1501 GCAAAAACAATTTGTAGAACAAAGTAGAAGCGGGCGCGATTTGGTTTAAATTTGCATGAAG 1560
DB 1501 GCAAAAACAATTTGTAGAACAAAGTAGAAGCGGGCGCGATTTGGTTTAAATTTGCATGAAG 1560

QY 1561 ACTGGGCGACAAACCAAGTCGATCGATCACTGTCTGAGCGTGGCAGATGAATACGATG 1620
DB 1561 ACTGGGCGACAAACCAAGTCGATCGATCACTGTCTGAGCGTGGCAGATGAATACGATG 1620

QY 1621 TGCAGTTTGTATCCACACCGCATCACTCAATGAGGCGAGTTATGTAGATGACACCTTAA 1680
DB 1621 TGCAGTTTGTATCCACACCGCATCACTCAATGAGGCGAGTTATGTAGATGACACCTTAA 1680

QY 1681 ATGCATGAACGGGCGCGCATCCATGCTTACCACTTGAAGGAGCGGTGGAGGACACT 1740
DB 1681 ATGCATGAACGGGCGCGCATCCATGCTTACCACTTGAAGGAGCGGTGGAGGACACT 1740

QY 1741 CACCTGATGTTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCCA 1800
DB 1741 CACCTGATGTTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCCA 1800

QY 1801 CTATTCCTTATACCAATTAATACGGTTGAGAAACACTTAGACATGCTCATGATGCCACC 1860
DB 1801 CTATTCCTTATACCAATTAATACGGTTGAGAAACACTTAGACATGCTCATGATGCCACC 1860

QY 1861 ACCTAGACAAACGCAATCCGCGAGGATTTACAAATTTTCTCAAGCGGTATCCGCGCCGGCT 1920
DB 1861 ACCTAGACAAACGCAATCCGCGAGGATTTACAAATTTTCTCAAGCGGTATCCGCGCCGGCT 1920

QY 1921 CTATCGCGCTGAGATGTGCTCCATGATATGGGTGTGATCCGATGACAAAGCTCGGATTT 1980
DB 1921 CTATCGCGCTGAGATGTGCTCCATGATATGGGTGTGATCCGATGACAAAGCTCGGATTT 1980

QY 1981 CGCAAGCAATGGGCGCTGAGCGGAGGATTTCTCGAACTTTGGCAGACTTGGCGATAAGA 2040
DB 1981 CGCAAGCAATGGGCGCTGAGCGGAGGATTTCTCGAACTTTGGCAGACTTGGCGATAAGA 2040

QY 2041 ATAAAAAGAAATTTGGTGAAGCTTCTGAAGATGGCAAGATGAACGATAATTTCCGCAATTA 2100
```

```
Db 2041 ATAAAAAGAAATTGGTAAGCTTCTCGAAGATGCCAAGATAACGATAATTTCCGCATT 2100
Qy |||||
Db 2101 AGCGTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGGGTGAGCAGATATA 2160
Db 2101 AGCGTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGGGTGAGCAGATATA 2160
Qy 2161 TCGGCTCTGTGGAGAGGCAAGATCCCGACTTGGTGTGGAATCCTGCCCTTTTGG 2220
Db 2161 TCGGCTCTGTGGAGAGGCAAGATCCCGACTTGGTGTGGAATCCTGCCCTTTTGG 2220
Qy 2221 GCCTAAAAACCAAAATCGTGATCAAGCGCGTATGGTGGTCTTCTCGAAATGGCGANT 2280
Db 2221 GCCTAAAAACCAAAATCGTGATCAAGCGCGTATGGTGGTCTTCTCGAAATGGCGANT 2280
Qy 2281 CTAAACCGCTCTGTGCCCATCTCCCAACCGGTTTATACCGCGAAATGTTGGGCGATCAG 2340
Db 2281 CTAAACCGCTCTGTGCCCATCTCCCAACCGGTTTATACCGCGAAATGTTGGGCGATCAG 2340
Qy 2341 GCAAGGCGAAATTGACACAGCATCACTTTGTTGTTTCCAAAGTCGCCCTATGAAATGGCG 2400
Db 2341 GCAAGGCGAAATTGACACAGCATCACTTTGTTGTTTCCAAAGTCGCCCTATGAAATGGCG 2400
Qy 2401 TGAAGAAAGAGCTGGGCTTAGAGCGCAAGTTCTTACCGGTCAAAAATCTGCCGTAAACATCA 2460
Db 2401 TGAAGAAAGAGCTGGGCTTAGAGCGCAAGTTCTTACCGGTCAAAAATCTGCCGTAAACATCA 2460
Qy 2461 CCAAGAAAGAGCTTCAAGTTCAACGACAAACCGCAAAAATCACCGTCGATCCGAAAAACCT 2520
Db 2461 CCAAGAAAGAGCTTCAAGTTCAACGACAAACCGCAAAAATCACCGTCGATCCGAAAAACCT 2520
Qy 2521 TCAGAGTCTTTGTAGATGGCAAACTCTGCACTTAAACCCACCTCGCAAGTGCTCTAG 2580
Db 2521 TCAGAGTCTTTGTAGATGGCAAACTCTGCACTTAAACCCACCTCGCAAGTGCTCTAG 2580
Qy 2581 CCAGCGCTACACTTCTTCTAGGCAAAATGCCCCCTTTGGGGCAGGTTATTTAGGA 2640
Db 2581 CCAGCGCTACACTTCTTCTAGGCAAAATGCCCCCTTTGGGGCAGGTTATTTAGGA 2640
Qy 2641 TCTTCATCAAAACGACCTGCAATCGGCTTTCGCGTGCAGATCGTGCCTTTAAACCAAC 2700
Db 2641 TCTTCATCAAAACGACCTGCAATCGGCTTTCGCGTGCAGATCGTGCCTTTAAACCAAC 2700
Qy 2701 TTTTCATCTTTAAGCAATGCCAATTTTAAATTAATTTAATTTCTTATTAATTAATATAT 2760
Db 2701 TTTTCATCTTTAAGCAATGCCAATTTTAAATTAATTTAATTTCTTATTAATTAATATAT 2760
Qy 2761 TATGCCCTCTATTTTAAAGGAGAAATATGCGTAGTCTTTGGTATGCTATGCGGGG 2820
Db 2761 TATGCCCTCTATTTTAAAGGAGAAATATGCGTAGTCTTTGGTATGCTATGCGGGG 2820
Qy 2821 TTGTTTGGTCTGGCGCAAGGATTTGAAACCCATCGCTCAAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTCTGGCGCAAGGATTTGAAACCCATCGCTCAAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883
```

```
RESULT 2
AX356683
LOCUS AX356683
DEFINITION Sequence 1 from Patent EP1176192.
ACCESSION AX356683
VERSION AX356683.1 GI:18674020
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
1
REFERENCE
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
```

```
JOURNAL Patent: EP 1176192-A 1 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
source
1. 2883
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
206. 886
/feature="unnamed protein product"
/transl_table=1
/protein_id="CAD23255.1"
/db_xref="GI:18674021"
/translation="MKLTPKQEKFLYYAGEVARKRKABGLKLNQPEALAIYSAIHM
DEARRGKRTVAQLMECHFLKQDRVMPGVGMVFDLGVTEATFPDGTGLVTVNMPLEP
DEHPKAGEVIFSGCDKDIENLAGKEVTEVNGPKSLHVGSHFIFFEANKLKFRE
KAYGKRLDIPEGNTLRI GAGQTRKVLQIPLGGSKKVIQNMGLVNNIADERHKHAKLDK
AKSHGFIK"
897. 2603
/feature="unnamed protein product"
/transl_table=1
/protein_id="CAD23256.1"
/db_xref="GI:18674022"
/translation="MKMKKQBYNTYGTGDKVRLGDTDLWAEVHDYTTTGEELKF
GAGKTIREGMQSNSPDENLIDLVTINAMI IDYTG IYKADI GIKNGKIHGIGKAGNKD
MQDGVSMVVGTEALAGEGMIITAGDISHTFLSPQOPTTALANGVITMFGGKT
PVDGTNATITPGKWNLMRLAAREYSMNVFLGKNGSSKQIQLVEVGAIGIFK
LHEDWGTTPSAIDHCLSVADRYDVQVCITHTDVTNEAGYVDDTLNANMGRALIHAYLEG
AGGSHSDVITMAGELNLPSTPTPIPTINTVAEHLDMLTCHILDKRILREDLQFS
QSRIRPSISIAADVLHDMGVJAMTSSDSQAMGRAGEVIRPTWTADADNKKFPGKLPED
KQNDNFRIIRKYSKITINPALTHGVSYIGSVEEGKIADLVNMPAFVGPVKIVIK
GMYVFSMGSDNASVPTPQPVYREMFHGHGKAKFDTSITFVSKVAYENGVEKLG
EROVLPVKNCRNITKDPFKNDKTAKITVDPKTFEVFDGKLCCTSKPTSQVPLAQRVT
FF"
ORIGIN
Query Match 100.0%; Score 2880.6; DB 2; Length 2883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 RGRAGATTTTCCARCACTTCAAGCACAATATGATCCTGTGTGTGGGTGGAATTCR 60
Db 1 RGRAGATTTTCCARCACTTCAAGCACAATATGATCCTGTGTGTGGGTGGAATTCR 60
Qy 61 ACTTGTTAACTCTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 120
Db 61 ACTTGTTAACTCTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 120
Qy 121 TTACTTATATTAATAAGTTAATAAAGTAACGAAATTAGGACTATAATCCCATTCGCTT 180
Db 121 TTACTTATATTAATAAGTTAATAAAGTAACGAAATTAGGACTATAATCCCATTCGCTT 180
Qy 181 TAAATTTAAACAAGAGTAATAGGTGAACCTACACCCAAAGAGCAAGAAAGTCTT 240
Db 181 TAAATTTAAACAAGAGTAATAGGTGAACCTACACCCAAAGAGCAAGAAAGTCTT 240
Qy 241 GTTATATATGCGGCGAAGTGCTAGAAAGCGCAAGCGACGAGGCTTAAGCTCAACCA 300
Db 241 GTTATATATGCGGCGAAGTGCTAGAAAGCGCAAGCGACGAGGCTTAAGCTCAACCA 300
Qy 301 ACCCGAAGCCATTCCTTACATTTAGTGCCTCATATTAATGACGAGCGCGCGGAAAAA 360
Db 301 ACCCGAAGCCATTCCTTACATTTAGTGCCTCATATTAATGACGAGCGCGCGGAAAAA 360
Qy 361 AACCGTTGCCAGCTTATGGAAGAGTGATGACACTTTTTGAAAAAAGATGAAGTATGCC 420
Db 361 AACCGTTGCCAGCTTATGGAAGAGTGATGACACTTTTTGAAAAAAGATGAAGTATGCC 420
Qy 421 CGGGGTGGTAAATATGTTTCCCGATCTAGGTGTAGAACCCACCTTTCTCTGATGGTACGAA 480
Db 421 CGGGGTGGTAAATATGTTTCCCGATCTAGGTGTAGAACCCACCTTTCTCTGATGGTACGAA 480
```

QY 481 ACTTGTAACGTGTAATTGGCCCATCGAACAGATGAGCACTTCAAAGCGGGGGAAGTGA 540
DB 481 ACTTTGTAACGTGTAATTGGCCCATCGAACAGATGAGCACTTCAAAGCGGGGGAAGTGA 540
QY 541 ATTTGGTTGGCATAAAGACATCGAGCTCAATGCGAGCAAGAGTAAACCGAACTTGAGGT 600
DB 541 ATTTGGTTGGCATAAAGACATCGAGCTCAATGCGAGCAAGAGTAAACCGAACTTGAGGT 600
QY 601 TACTAATAAGAGGCGCTAAATCCTTGTCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAA 660
DB 601 TACTAATAAGAGGCGCTAAATCCTTGTCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAA 660
QY 661 CAAGGCACTAAAATTCGATCGTGA AAAAGCCTATGCGAAAACGCTAGATATTCCTCTGG 720
DB 661 CAAGGCACTAAAATTCGATCGTGA AAAAGCCTATGCGAAAACGCTAGATATTCCTCTGG 720
QY 721 CAACAGCTACGCAATTGGGGGAGGACAAACCGCAAGTGCAGTTGATTCCTCTGGTGG 780
DB 721 CAACAGCTACGCAATTGGGGGAGGACAAACCGCAAGTGCAGTTGATTCCTCTGGTGG 780
QY 781 CAGTAAAAAAGTGAATTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
DB 781 CAGTAAAAAAGTGAATTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
QY 841 ACATAAAGCGCTTGACAAGGCGAAATCTCACGGAATTAACGATTAAGTAAAGGAGACTCCCATGA 900
DB 841 ACATAAAGCGCTTGACAAGGCGAAATCTCACGGAATTAACGATTAAGTAAAGGAGACTCCCATGA 900
QY 901 AAATGMAAAACAGAAATGTAAATACCTACGGACCCACCAAGGCGATAAAGTGCCT 960
DB 901 AAATGMAAAACAGAAATGTAAATACCTACGGACCCACCAAGGCGATAAAGTGCCT 960
QY 961 TAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTATGGCGAAGAAC 1020
DB 961 TAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTATGGCGAAGAAC 1020
QY 1021 TTAATTTGGCGCGGTA AAAACTATCCGTGAGGTTATGGGTAGAGCAATAGCCCTGATG 1080
DB 1021 TTAATTTGGCGCGGTA AAAACTATCCGTGAGGTTATGGGTAGAGCAATAGCCCTGATG 1080
QY 1081 AAAACACCTAGATTAGTCACTAATACCGGATGATTTATCGACTACACCGGGATTACA 1140
DB 1081 AAAACACCTAGATTAGTCACTAATACCGGATGATTTATCGACTACACCGGGATTACA 1140
QY 1141 AAGCCGATTTGGGATTA AAAACGCAAAATCCATGGCAATGGCAAGCAGGAAACAAG 1200
DB 1141 AAGCCGATTTGGGATTA AAAACGCAAAATCCATGGCAATGGCAAGCAGGAAACAAG 1200
QY 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTGTGGGTGTGGGCAAGAAGCACTAGCAG 1260
DB 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTGTGGGTGTGGGCAAGAAGCACTAGCAG 1260
QY 1261 GGGAAAGTATGATTAATACCGCTGGGGGAATCGATTACACACCACTTCTTCTCCAC 1320
DB 1261 GGGAAAGTATGATTAATACCGCTGGGGGAATCGATTACACACCACTTCTTCTCCAC 1320
QY 1321 AACAAATTCCTACCGCTCTAGCCATGCGTTTACAAACCATGTTTGGAGCGGCAAGGTC 1380
DB 1321 AACAAATTCCTACCGCTCTAGCCATGCGTTTACAAACCATGTTTGGAGCGGCAAGGTC 1380
QY 1381 CTGTAGATGGCAAGATCGACTACTATCACTCCGGGCAAAATGGAACTTGACACCGCATGT 1440
DB 1381 CTGTAGATGGCAAGATCGACTACTATCACTCCGGGCAAAATGGAACTTGACACCGCATGT 1440
QY 1441 TGGCGGCAAGCAAGATTAATCTAATGATGGGTTTTTGGGCAAAAGCAATAGCTCTA 1500
DB 1441 TGGCGGCAAGCAAGATTAATCTAATGATGGGTTTTTGGGCAAAAGCAATAGCTCTA 1500
QY 1501 GCMAAAACAACTTGTAGAACAGTAGAGCGGCGGATTTGGTTTAAATTCGATGAG 1560
DB 1501 GCMAAAACAACTTGTAGAACAGTAGAGCGGCGGATTTGGTTTAAATTCGATGAG 1560
QY 1561 ACTGGGGGCAACACCAAGTGGATCGATCACTGTGTGAGCGTGGCAGATGAATACGATG 1620

DB 1561 ACTGGGGGCAACACCAAGTGGATCGATCACTGCTTGAGCGTGGCAGATGAATACGATG 1620
QY 1621 TGCAGATTTGTATCCACACCGGATCAGTCAATGAGGACGTTATGTAGATGACACCCCTAA 1680
DB 1621 TGCAGATTTGTATCCACACCGGATCAGTCAATGAGGACGTTATGTAGATGACACCCCTAA 1680
QY 1681 ATGCAATGAACGGGCGGCGCATCCATGCTTACCAATTCGAGGAGCGGTGGAGGACCT 1740
DB 1681 ATGCAATGAACGGGCGGCGCATCCATGCTTACCAATTCGAGGAGCGGTGGAGGACCT 1740
QY 1741 CACCTGATTTATCACATGCGAGCGGAGCTCAATATTTCTACCTCTCTCAACACCCCA 1800
DB 1741 CACCTGATTTATCACATGCGAGCGGAGCTCAATATTTCTACCTCTCTCAACACCCCA 1800
QY 1801 CTATTCCTTATACCAATTAATACGTTTGCAGAACACTTTAGACATGCTCATGACATGCCACC 1860
DB 1801 CTATTCCTTATACCAATTAATACGTTTGCAGAACACTTTAGACATGCTCATGACATGCCACC 1860
QY 1861 ACCTAGACAAACGCAATCGCGAGGATTTACAATTTTCTCAAAAGCGGTATCCGCCCGGT 1920
DB 1861 ACCTAGACAAACGCAATCGCGAGGATTTACAATTTTCTCAAAAGCGGTATCCGCCCGGT 1920
QY 1921 CTATCGCGCTGAAGATGTCTTCCATGATATGCGTGTGATCGCGATGACAAGCTCGATT 1980
DB 1921 CTATCGCGCTGAAGATGTCTTCCATGATATGCGTGTGATCGCGATGACAAGCTCGATT 1980
QY 1981 CGCAAGCAATGGGCGTGCAGGCGAAGTATTCCTCGAACTTGGGACACTTGGCGATAAGA 2040
DB 1981 CGCAAGCAATGGGCGTGCAGGCGAAGTATTCCTCGAACTTGGGACACTTGGCGATAAGA 2040
QY 2041 ATAAAAAGAAATTTGGTAACTTCTGAAAGATGGCAAGATAACGATAATTTCCGCAATTA 2100
DB 2041 ATAAAAAGAAATTTGGTAACTTCTGAAAGATGGCAAGATAACGATAATTTCCGCAATTA 2100
QY 2101 AGCGCTATCTCCTCAAAATACATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATA 2160
DB 2101 AGCGCTATCTCCTCAAAATACATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATA 2160
QY 2161 TCGGCTCTGTGGAAGGCGAAGATCGCGACTTGGTGTGTGGAATCTGCTCTTTTTCG 2220
DB 2161 TCGGCTCTGTGGAAGGCGAAGATCGCGACTTGGTGTGTGGAATCTGCTCTTTTTCG 2220
QY 2221 GCGTAAAAACCCAAATTCGTGATCAAAAGCGGTATGGTGTCTTCTGAAATGGGCGATT 2280
DB 2221 GCGTAAAAACCCAAATTCGTGATCAAAAGCGGTATGGTGTCTTCTGAAATGGGCGATT 2280
QY 2281 CTAAACGCTGTGTGCCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCACTACG 2340
DB 2281 CTAAACGCTGTGTGCCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCACTACG 2340
QY 2341 GCAAGCGAAATTTGACACCGCATCATTTTGTTCCTCAAAAGTCCGCTTGAATAATGGCG 2400
DB 2341 GCAAGCGAAATTTGACACCGCATCATTTTGTTCCTCAAAAGTCCGCTTGAATAATGGCG 2400
QY 2401 TGAAGAAAAAGCTGGGCTTAGAGCGCAAGTCTACCGGTCAAAAACTCGCGTAAACATCA 2460
DB 2401 TGAAGAAAAAGCTGGGCTTAGAGCGCAAGTCTACCGGTCAAAAACTCGCGTAAACATCA 2460
QY 2461 CCAAGAAAGACTTCAAGTTCAACGACAAACCGCAAAATCACCGTCGATCCGAAACCT 2520
DB 2461 CCAAGAAAGACTTCAAGTTCAACGACAAACCGCAAAATCACCGTCGATCCGAAACCT 2520
QY 2521 TCGAGCTCTTGTAGATGGCAAACTCTGCACTTAAACCCACCTCGCAAGTCCCTCTAG 2580
DB 2521 TCGAGCTCTTGTAGATGGCAAACTCTGCACTTAAACCCACCTCGCAAGTCCCTCTAG 2580
QY 2581 CCCAGCGCTACACTTTCTTAGGCACAATGCCCCCTTTGGGGGCGAGGTTATTTTAGGAA 2640
DB 2581 CCCAGCGCTACACTTTCTTAGGCACAATGCCCCCTTTGGGGGCGAGGTTATTTTAGGAA 2640
QY 2641 TCTTCATCAACGCACTCGTCAATCGGTCTTGGGTGTGCGATCGTGTCTTTTAAACAAC 2700

```

Db 2641 TCCTTCATCAAAACGACCTCGCAATCGTCTTGGTGTGCGATCGTGTGCTTAAACACAC 2700
Qy 2701 TTTTCATCTTTAAGCAATCGCCATTTTAAATTAATTTAATCTTATAATTAATATATAT 2760
Db 2701 TTTTCATCTTTAAGCAATCGCCATTTTAAATTAATTTAATCTTATAATTAATATATAT 2760
Qy 2761 TATGCCCTCATTTTAAAGGAGAAATATGCTAGTCTTTGGTATTTGCTATGTGGGG 2820
Db 2761 TATGCCCTCATTTTAAAGGAGAAATATGCTAGTCTTTGGTATTTGCTATGTGGGG 2820
Qy 2821 TTGTTTGGTGTGGCGCAAAAGGATTTGAAACCCATCGCTCAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTGTGGCGCAAAAGGATTTGAAACCCATCGCTCAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 3
BD185306
LOCUS Helicobacter felis vaccine. 2452 bp DNA linear PAT 17-JUN-2003
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185306
VERSION BD185306.1 GI:31877506
KEYWORDS JP 2002355054-A/5.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2452)
AUTHORS Kusters,J.G. and Cattolli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: JP 2002355054-A 5 10-DEC-2002;
AKZO NOBEL NV
COMMENT OS Helicobacter felis
PN JP 2002355054-A/5
PD 10-DEC-2002
PF 16-JUL-2001 JP 2001214711
PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLLI
PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
A61K39/23,
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
PC C12N1/19,
PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC
53,
PC G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1/01),(C12Q1/68, PC
C12R1/01),
PC C12N15/00,C12N5/00,A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (48)..(728)
FT CDS (739)..(2445).
FEATURES
source 1..2452
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"

Query Match 77.3%; Score 2228; DB 2; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 159 AGGACTATAATCCCATGCTTTAAATTTAAACCAAGAGGTAATAGTGAACCTCACAC 218
Db 1 AGGACTATAATCCCATGCTTTAAATTTAAACCAAGAGGTAATAGTGAACCTCACAC 60
Qy 219 CCAAAGAGCAAGAAAGTTCTTGTATTATTATGCGGGCGAAGTGGCTAGAAAGCGCAAAG 278

```

```

Db 61 CCAAGAGCAAGAAAGTTCTTGTATTATATGCGGGCGAAGTGGCTAGAAAGCGCAAAG 120
Qy 279 CAGAGGGCTTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGTGCCCATATTATGG 338
Db 121 CAGAGGGCTTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGTGCCCATATTATGG 180
Qy 339 ACGAAGCGCGCGTGGAAAAAAAACCCGTTGCCAGCTTATGGAAGAGTGCAATGCACTTTT 398
Db 181 ACGAGCGCGTGTGCGCAAAAAAACCCGTTGCCGAACCTTATGGAAGAGTGATGCACTTTT 240
Qy 399 TCAAAAAGATCAAGTAATGCGCGGGTGGTAATATGTTCCCGATCTAGGTGTAGAAG 458
Db 241 TGAaaaaagacacaggtgatgccccggggggaatatggtccctgattttggggcggtgaag 300
Qy 459 CCACCTTTCTCATGTACGAAACTTGTAACTGTGAAATTTGGCCCATCGAACACAGATGAGC 518
Db 301 CCACCTTTCCCGATGGCACCAAACTCGTAACTGTGAAATTTGGCCCATCGAACCTGATGAAC 360
Qy 519 ACTTCAAGCGGGCGAAGTGAATAATTTGTTGCGATAAAGACATCGAGCTCAATGCAAGCA 578
Db 361 ACTTTAAGCGGGTGAAGTGAATAATTTGGCTGTGATTAAGACATTTGAACCTCAACGCAAGTA 420
Qy 579 AAGAAGTAACCGAACTTGAGGTACTTAATGAAGGGCTTAAATCTTTCATGTTGGGTAGCC 638
Db 421 AGGAAGTTTACCGAACTAGAAGTTACTAACGAAGGACCTTAAATCTTTCATGTTGGGTAGCC 480
Qy 639 ATTTCCACTTTCTCAAGCTAACCAAGGCACTAAATTTGATCGTGAATAAAGCCCTATGGCA 698
Db 481 ATTTCCACTTTCTTGAAGCCAAAGCAATGGAATTCGATCGGGAATAAGCCCTATGGCA 540
Qy 699 AACGCTAGATATTCCCTCTGCGCAACACGCTACGCAATTTGGGGCAGGACAAACCCGCAAAG 758
Db 541 AAGCCTAGATATTCCCTCTGCGCAACACACTAGGCAATTTGGGGCAGGACAAACCCGTAAG 600
Qy 759 TGCAGTTGATTCTCTTGGTGGCAGTAAATAAGTGAATTTGGCATGAACGGGCTTTGTGAATA 818
Db 601 TGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAGTGAATTTGGCATGAACGGGCTTTGTGAATA 660
Qy 819 ACATCGCGGATGAAGCCATAAACAATAAGCGCTTTCACAAGGCGGAATCTCACGGATTTA 878
Db 661 ATATTGCGGACGACGCGCATAAACAACAAGCGCTAGACAAAGCAAAATCTCACGGATTTA 720
Qy 879 TCAAGTAAGGAGACTCCCATGAAATGAAAAACAAGAAATATGTAATACCTACGAGACCC 938
Db 721 TCAGTAAGGAGACTCCCATGAAATGAAAAACAAGAGTATGTAATACCTACGAGACCC 780
Qy 939 ACCAAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTTGGGCGAGAAGTAGAACATGAC 998
Db 781 ACCACAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTTGGGCGAGAAGTAGAACATGAC 840
Qy 999 TATACCCTATGGGGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTGAGGGTATG 1058
Db 841 TATACCCTATGGGGAAGAACTCAAAATTCGGTGCAGGTAAAACTATCCGTGAGGGTATG 900
Qy 1059 GGTACAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTATCATCACTAACGCGATGATT 1118
Db 901 GGTACAGCAATAGCCCGAGATGAAAAACACCTTAGATTTAGTATCATCAACACGCGATGATT 960
Qy 1119 ATCGACTACACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCCATGGC 1178
Db 961 ATTGACTACACCGGGATTTACAAAGCCGACATTTGGCATTTAAAAATGGCAAAATCCCATGGC 1020
Qy 1179 ATTGGCAGGCGAGAAACAAGACATGCAAGATGGCGGTAAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGGCAGGCGAGAAACAAGACATGCAAGATGGCGGTAAAGCCCTCATATGTCGTGGGT 1080
Qy 1239 GTGGGCGACAGAGCACTAGCAGGGGAGAGTATGATTATTACCGCTGGGGGAATCCGATTCA 1298
Db 1081 GTGGGCGACAGAGCACTAGCAGGGGAGAGTATGATTATTACCGCTGGGGGAATCCGATTCA 1140
Qy 1299 CACACCCACTTCTCTTCTTCTCCACAACAATTCCTACCGCTTAGCCCAATGGCGCTTACAAC 1358

```

Db 1141 CACACCCACTTCCTCTCCACAACAATTCCTACCGCTCTAGCCAATGGCGTTACAACA 1200
Qy 1359 ATGTTTGGAGGCGGACAGGCTCTCTAGATGGCAAGATGGAGCTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGTGGACAGGCGCCGTAGATGGCACGAATGGAGCTACCATCACTCCGGGC 1260
Qy 1419 AAATGGAACTTTCACCGCATGTTGGCGGCAGCAGAGAGATTTCTATGAATGTGGGCTTT 1478
Db 1261 AAATGGAACTTTCACCGCATGTTGGCGGCAGCAGAGAGATTTCTATGAATGTGGGCTTT 1320
Qy 1479 TTGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAGTAGTAGAGCGGGCGG 1538
Db 1321 TTGGCAAGGCAATAGCTCTAGTAAAAAAACAATTTGTAGAACAGTAGTAGAGCGGGCGG 1380
Qy 1539 ATTGGTTTAAATTCATGAAGAATGGGGCAACAACCAAGTGGCATCGATCACTGCTTG 1598
Db 1381 ATTGGTTTAAATTCATGAAGACTGGGGCAACAATCCAAAGTGGCATCGATCACTGCTTG 1440
Qy 1599 AGCGTGGCAGATGAATAGATGTGCAAGTTTGTATCCACCGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTGGCAGATGAATAGATGTGCAAGTTTGTATACACCGATACAGTCAATGAGGCA 1500
Qy 1659 GGTATGTAGATGACACCTCAATGAATGAACGGGGCGGCATCCATCGCTTACCACATT 1718
Db 1501 GGTATGTAGATGACACCTCAATGAATGAACGGGGCGGCATCCATCGCTTACCACATT 1560
Qy 1719 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGCGAGTGAATATT 1620
Qy 1779 CTACGCTCTCCACACCGCCCTACTATCCCTATACCATTAATACGGTTGCGAGAACACTTA 1838
Db 1621 CTACGCTCTCCACACCGCCCTACTATCCCTATACCATTAATACGGTTGCGAGAACACTTA 1680
Qy 1839 GACATGCTCATGACATGCCACCATCAGACCAACCGCATCCGCGAGGATTACAAATTTCT 1898
Db 1681 GACATGCTTATGACTGCCACCATCAGACCAACCGCATCCGCGAGGATTCCAAATTTCT 1740
Qy 1899 CAAAGCCGTATCCGCGCGGCTCTATCGCGGTGAAGATGTGCTCCATGATATGGGTGTG 1958
Db 1741 CAAAGCCGTATCCGCGCGGCTCTATCGCGGTGAAGATGTGCTCCATGATATCGGTGTG 1800
Qy 1959 ATCGGATGACAAGCTCGGATTCGGAAGCAATGGCGGTGCGAGGCAAGTGAATTCCTGA 2018
Db 1801 ATCGGATGACAAGTTCGGATTCGGAAGCAATGGCGGTGCGAGGCAAGTGAATTCCTGA 1860
Qy 2019 ACTTGGCAGCTCGGATGAAGATTAAGAAGAAATTTGGTGAAGCTTCCTGAAGATGGCAA 2078
Db 1861 ACTTGGCAGCTCGGATGAAGATTAAGAAGAAATTTGGTGAAGCTTCCTGAAGATGGTGA 1920
Qy 2079 GATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAACTTCGGATCAACGCTATATCTCCAAATACACCAATTAATCCCGCTTTG 1980
Qy 2139 ACCCAACGGCGTGAAGGATATATCGGCTCTGGAAGAGGGCAAGATCGCGACTTGGTG 2198
Db 1981 ACCCAACGGCGTGAAGGATATATCGGCTCTGGAAGAGGGCAAGATCGCGACTTGGTG 2040
Qy 2199 GTGTGGAATCTCGGCTTTTGGCGTAAACCCAAATCGTGATCAAGGCGGTATGTTG 2258
Db 2041 GTGTGGAATCTCGGCTTTTGGCGTAAACCCAAATCGTGATCAAGGCGGTATGTTG 2100
Qy 2259 GTCTTCTCTGAATGGCGGATTTCAACCGCTCTGTGCCCATCTCCCAACCGGTTTATTAC 2318
Db 2101 GTGTTCTCTGAATGGCGGATTTCTAATGCGTCTGTGCCCATCTCTCAGCGGTTTATTAC 2160
Qy 2319 CGCGAAATGTTTGGGCATCAGCGCAAGGCGAAATTTGACACAGCATCACTTTTGTTC 2378
Db 2161 CGCGAAATGTTTGGGCATCAGCGCAAGGCGAAATTTGACACAGCATCACTTTTGTTC 2220
Qy 2379 AAAGTCGCTATGAATGGCGTGAAGAAAGAGCTGGGCTTAGAGCGCGCAAGTCTACGG 2438
Db 2221 AAAGTCGCTATGAATGGTGTGAAGAAAGAACTAGGTTTAGAGCGCAAGGTGCTCC 2280

Qy 2439 GTCAAAACTGCGTAACATCAACGAAGAAGACTTCAAGTTCAACGAACAAACGCGAAA 2498
Db 2281 GTGAAAAACTGCGTAACATCAACGAAGAAGACTTCAAGTTCAACGAACAAACGCGAAA 2340
Qy 2499 ATCACCGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGGCAACTCTGCACCTTAA 2558
Db 2341 ATCACCGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGGCAACTCTGCACCTTAA 2400
Qy 2559 CCCACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTAGGCACAAT 2610
Db 2401 CCCACTCTGAAGTGCCTCTAGCCACGCTACACTTTCTTAGGCATAAT 2452

RESULT 4
AX356695
LOCUS AX356695 2452 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 13 from Patent EP1176192.
ACCESSION AX356695
VERSION AX356695.1 GI:18674032
KEYWORDS
SOURCE
ORGANISM
Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.

REFERENCE
1
AUTHORS Kusters, J.G. and Cottoli, G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: EP 1176192-A 13 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
Location/Qualifiers
1..2452
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
48..728
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23263.1"
/db_xref="GI:18674033"
/translation="MKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIM
DEARRKKTVAELMECHFLKKDRVMPGQNMVPLDGVETFPDGTKLVTWNPLEP
DEHFKAGVFCGDKDI ELNAGKRYTELVNEGPKSLHVGSHFHF FEANKLKP RE
KAYGRLDI PSGNTLRI GAGTRKVL I PLGGSKVI GMNLVNNI ADERHKH KALDK
AKSHGFIR"
739..2445
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23264.1"
/db_xref="GI:18674034"
/translation="MKMKQEVNTYGPPTGDKVRLGDTDLNAEVEHDYTTGEEELKF
GAGTIRGMSGNSPDENLTLVITNAMIIDYTGILKADIGIKNGIKHIGKAGNKD
MDQVSPHMVYGTGEALAGBEMIITAGIDSHFLSPQPTALANGVTTFPGGT
GPDGTNATITTPGKNLHRLMRAAEYSMNQFLGKNSKSKQI LVEGAEAGI GFK
LHEDGTTPSAIDHCLSVADYVQVCIHTDVTNAGYVDDTLNMGRAI HAYHIEG
AGGHSPIVITMAGEVNILPSSTTPTIPYITVAEHLDMLTCHLDRK IREDLOFS
QSRIRPGSIAEDVLHDIGVIAMTSSDSQAMGRAGEV IPRWTQADKNKKEPKLPED
GADNDFRIKRIKYITINPALTGVSEYIGSEBGKIADLVVNNPAPFFGVKPKIVIK
GMVYFSGMSGNSASVPTPPVYIREMFGHHGAKFPDTSITTFVSKVAYENGVBKILGL
ERKVLVPVKNCRNITKKDFKENDKATKITVDPKTFEYFVDGKLTCKTSKPTSEVPLAQRYT
FP"

ORIGIN

Query Match 77.3%; Score 2228; DB 2; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 159 AGGACTAATCCCATTCGCTTTAAATTTAAACAAGGAGTAATAGGTGAACACTCACAC 218
Db 1 AGGACTAATCCCATTCGCTTTAAATTTAAACAAGGAGTAATAGGTGAACACTCACAC 60

QY 219 CCAGAGGAGGAGGAGTTCTTGTATATATGTCGGGCGAAGTGGCTAGAAAGCGCAAG 278
DB 61 CCAAGAGCAGAGAGAGTTCTTGTATATATGTCGGGCGAAGTGGCTAGAAAGCGCAAG 120
QY 279 CAGAGGGCTTAAAGCTCAACCAACCGGAGGCAATTTGCTTACATTAAGTGGCCCATATATG 338
DB 121 CAGAGGGCTTAAAGCTCAACCAACCGGAGGCAATTTGCTTACATTAAGTGGCCCATATATG 180
QY 339 ACAGAGGGCGGTGGGAAAAAACCCTGTCGCCAGCTTATGGAAGAGTGCATGCACTTTT 398
DB 181 ACAGAGGGCGGTGGGCAAAAAACCCTGTCGGAACTTATGGAAGAGTGTATGCACTTTT 240
QY 399 TGAAGAGAGTGAAGTAATGCCGGGGTGGTATATGTTCCCGATCTAGTGTGAGAG 458
DB 241 TGAAGAGAGTGAAGTATGTCGGGGTGGGAAATATGTTCCCTGATTTGGGCGTGGAG 300
QY 459 CCACCTTTCTGATGTCGACCAACTCTGTAAGTGTGAATTTGGCCCATCGAACAGATGAGC 518
DB 301 CCACCTTTCCCGATGGGACCAACTCGTGAATTTGGCCCATCGAACCTGATGAGC 360
QY 519 ACTTCAAAGCGGGGAGAGTGAATTTGGTTGCGATAAGACATCGAGCTCAATCGAGCA 578
DB 361 ACTTTAAGCGGGTGAAGTGAATTTGGCTGTGATAAGACATTTGAACTCAACGCAAGTA 420
QY 579 AAGAGTAAACCGAATTTGAGGTTACTAATGAGGGCCCTAAATCCTTGCATGTGGGTAGCC 638
DB 421 AGGAAGTTACCGAATTAAGAGTTACTAAGCAAGGACCTAAATCCTTGCATGTGGGTAGCC 480
QY 639 ATTTCCACTTTCTTGAAGCTTAAAGGACCTTAAATTTGCGTGTGAAGAAAGCCCTATGCA 698
DB 481 ATTTCCACTTTCTTGAAGCCAAAGACATTTGAATTTGCGTGTGAAGAAAGCCCTATGCA 540
QY 699 AAGCCCTAGATATTTCCCTCTGGCAACACGCTACGATTTGGGGCAGGACAAACCCGCAAG 758
DB 541 AAGCCCTAGATATTTCCCTCTGGCAACACATACGCAATTTGGGGCAGGACAAACCCGTAAG 600
QY 759 TGCAAGTTGATTTCTTGGTGGCAGTAAAAAAGAGTGAATGGCATGAACGGGCTTGTGAATA 818
DB 601 TGCAAGTTAAATCCCTCTTGGCGGTAGTAAAAAAGTGAATGGCATGAACGGGCTTGTGAATA 660
QY 819 ACATCGGGATGAACGCATAAACATAAAGCGCTTGACAGGCGAATCTCACGGATTTA 878
DB 661 ATATTGGGAGCAGACGCCATAAACACAAAGCGCTTAGACAAAGCAAAATCTCACGGATTTA 720
QY 879 TCAAGTAAGGAGACTCCCATGAAAAATGAAAAAACAAGATATGTAATATCTACCGGACCC 938
DB 721 TCAAGTAAGGAGACTCCCATGAAAAATGAAAAAACAAGATATGTAATATCTACCGGACCC 780
QY 939 ACCAAGGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAACATGAC 998
DB 781 ACCAAGGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAACATGAC 840
QY 999 TATACCACTATGGCGAGAACCTTAAATTTGGGCGGGTAAACTATCCGTGAGGGTATG 1058
DB 841 TATACCACTATGGCGAGAACCTTAAATTTGGGCGAGGTAAAACTATCCGTGAGGGTATG 900
QY 1059 GGTTCAGAGCAATAGCTGATGAAAAACACCTAGATTTAGTTCATCACTAAACGGCATGAT 1118
DB 901 GGTTCAGAGCAATAGCTGATGAAAAACACCTAGATTTAGTTCATCACTAAACGGCATGAT 960
QY 1119 ATCGACTACACCGGGATTTACAAAGCGGACATTTGGGATTAATAACCGCAAAATCCATGGC 1178
DB 961 ATTTGACTACACCGGGATTTACAAAGCGGACATTTGGCATTAATAATGCAAAATCCATGGC 1020
QY 1179 ATTTGGCAGGCGAGAAACAAAGGACATGCAAGATGGCGTAAGCCCTCATATGTGGTGGGT 1238
DB 1021 ATTTGGCAAGGCGAGAAACAAAGGACATGCAAGATGGCGTAAGCCCTCATATGTGGTGGGT 1080
QY 1239 GTGGGACAGAACCATAGCAGGGGAGGTATGATTTATACCGCTGGGGGAGTTCGATTTCA 1298
DB 1081 GTGGGACAGAACCATAGCAGGGGAGGTATGATTTATACCGCTGGGGGAGTTCGATTTCA 1140
QY 1299 CACACCCACTTCTCTTCCACAAATTTCCCTACCGCTCTAGCCATGGCGTTACAAACC 1358

DB 1141 CACACCCACTTCTCTCTCCACAAATTTCCCTACCGCTCTAGCCATGGCGTTACAAACA 1200
QY 1359 ATGTTTGGAGGCGGCACAGGTCCTGTAGATGGCAAGAAATGCGACTACTATCACTCCGGGC 1418
DB 1201 ATGTTTGGCGGTGGCACAGGCCCTGTAGATGGCAAGAAATGCGACTACTATCACTCCGGGC 1260
QY 1419 AAATGGAACTTGCACCGCATGTTGGCGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTT 1478
DB 1261 AAATGGAACTTGCACCGCATGTTGGCGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTT 1320
QY 1479 TTGGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAAACAAGTAGAAGCGGGCGG 1538
DB 1321 TTGGGCAAGGCAATAGCTCTAGTAAAAAACAATTTGTAGAAACAAGTAGAAGCGGGCGG 1380
QY 1539 ATTGTTTTAAATTTGATGAAGACTGGGGCAACAACCAAGTGCATCGATCACTGCTTG 1598
DB 1381 ATTGTTTTAAATTTGATGAAGACTGGGGCAACAATTTCAAGTGCATCGATCACTGCTTG 1440
QY 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTGTGTATCCACCGATACAGTCAATGAGGCA 1658
DB 1441 AGCGTAGCAGATGAATACGATGTGCAAGTTGTGTATCCACCGATACAGTCAATGAGGCA 1500
QY 1659 GGTATGTAGATGACACCTTAAATGCAATGAACGGGCGCCATCCATGCTTACCACTT 1718
DB 1501 GGTATGTAGATGACACCTTAAATGCAATGAACGGGCGCCATCCATGCTTACCACTT 1560
QY 1719 GAGGAGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGCGAGCTCAATATT 1778
DB 1561 GAGGAGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGCGAGTGAATATT 1620
QY 1779 CTACCTCTCCACCAACCCCTACTATCCCTATACCAATTAACGGTTGAGAACTTTCT 1838
DB 1621 CTACCTCTCCACCAACCCCTACTATCCCTATACCAATTAACGGTTGAGAACTTTCT 1680
QY 1839 GACATGCTCATGACATGCGCACCCCTAGACAAACCGCATCCGCGAGGATTTTACAAATTTCT 1898
DB 1681 GACATGCTTATGACCTGCGCACCCCTAGATAAACCGCATCCGCGAGGATCTTCCAAATTTCT 1740
QY 1899 CAAGCGGTATCCGCGCGGCTCTATCCGCGGTGAAGATGTGCTCCATGATATGGGTGTG 1958
DB 1741 CAAGCGGTATCCGCGCGGCTCTATCCGCGGTGAAGATGTGCTCCATGATATGGGTGTG 1800
QY 1959 ATCCGATGACAAAGCTCGGATTTGCAAGCAATTTGGGCGGTGAGGCGAGTGTCTCTCGA 2018
DB 1801 ATCCGATGACAAAGTTTCGATTTGCAAGCAATTTGGGCGGCTGGGAGTGTCTCTCTAGA 1860
QY 2019 ACTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTAAAGCTTCTCTGAAGTGCATAA 2078
DB 1861 ACTTGGCAAACTGCGAGACAGATAAAGAAATTTGGTAAAGCTTCTCTGAAGTGTGCA 1920
QY 2079 GATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTG 2138
DB 1921 GATATGACAACTTCCGATCAAGCGCTATATCTCCAAATACCAATTAATCCCGCTTTG 1980
QY 2139 ACCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGGGGCAAGATCCCGACTTTGGTG 2198
DB 1981 ACCCATGGCTGAGCGAGTATATCGGCTCTGTGGAAGGGGCAAGATCCCGACTTTGGTG 2040
QY 2199 GTGTGGAAATCTGCTTTTTTTGGCGTAAAAACCCAAATCTGTGATCAAGCGGTATGGTG 2258
DB 2041 GTGTGGAAATCTGCTTTTTTTGGCGTAAAAACCCAAATCTGTGATCAAGCGGTATGGTG 2100
QY 2259 GTCTTCTCTGAATGGGCGATTTCTAACCGCTGTGGCCACTCCCAACCGGTTTATTAC 2318
DB 2101 GTGTCTCTGAATGGGCGATTTCTAAATCGCTGTGGCCACTCTCTCAGCGGTTTATTAC 2160
QY 2319 CGCGAAATGTTTGGGCATCAGCGCAAGCGGAAATTTTGACACAGCATCACTTTGTTTTC 2378
DB 2161 CGCGAAATGTTTGGGCATCAGCGCAAGCGGAAATTTTGACACAGCATCACTTTGTTTTC 2220
QY 2379 AAAGTCGCTATGAAATGGCGTGAAGAAAGAGTGGGCTTAGAGCGCAAGTTCCTACCG 2438

Db	2221	AAAGTCGCCTATGAAATAATGGTGTGAAAGAAAAACAATAGTGTAGAGCGCAAGGTGCTCCCC	2288
Qy	2439	GTCAAAAACATGCCGTAAATCATCACCAAGAAAGACTTCAAGTTTCAACGACAAAACGGCAAAA	2498
Db	2281	GTGAAAAATCGCGTAACATCACCAAGAGACTTCAAGTTTCAACGACAAAACGGCAAAA	2340
Qy	2499	ATCACCGTCGATCCGAAAAACCTTCGAGGCTTGTGTAGATGGCAAACTCTGCACCTCTAAA	2558
Db	2341	ATCACCGTCGATCCGAAAAACCTTCGAGGCTTGTGTAGATGGCAAACTCTGCACCTCTAAA	2400
Qy	2559	CCCACTCGCAAGTGCCTCTAGCCACGCGTACACTTTCTTCTAGGCACAAAT	2610
Db	2401	CCCACTCTGAAGTGCTCTAGCCCAACGCTACACTTTCTTCTAGGCATAAT	2452
RESULT 5			
LOCUS	BD185303	2405 bp	DNA linear PAT 17-JUN-2003
DEFINITION	Helicobacter felis vaccine.		
ACCESSION	BD185303		
VERSION	BD185303.1	GI:31877503	
KEYWORDS	JP 2002355054-A/2.		
SOURCE	Helicobacter felis		
ORGANISM	Helicobacter felis		
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
AUTHORS	1 (bases 1 to 2405)		
TITLE	Kusters, J.G. and Cattoli, G.		
JOURNAL	Helicobacter felis vaccine		
COMMENT	Patent: JP 2002355054-A 2 10-DEC-2002; AKZO NOBEL NV		
COMMENT	OS	Helicobacter felis	
	PN	JP 2002355054-A/2	
	PD	10-DEC-2002	
	PF	16-JUL-2001 JP 2001214711	
	PR	17-JUL-2000 EP 00202585.8	
	PI	JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI	
	PC		
	C12N1/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC		
	A61K39/23,		
	PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15, PC C12N1/19,		
COMMENT	PC C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, 53,		
	PC G01N33/53, G01N33/566, G01N33/569/(C12N9/80, C12R1/01), (C12Q1/68, PC C12R1/01),		
	PC C12N15/00, C12N5/00, A61K37/02		
	CC Helicobacter felis vaccine		
	PH Key Location/Qualifiers		
	FT CDS (1) (681)		
	FT CDS (692) (2398).		
	Location/Qualifiers		
	1. 2405		
	/organisms="Helicobacter felis"		
FEATURES	/mol_type="genomic DNA"		
	/db_xref="taxon:214"		
	source		
	Query Match 74.9%; Score 2158.6; DB 2; Length 2405;		
	Best Local Similarity 93.6%; Pred. No. 0;		
	Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;		
Qy	206	GTGAAACTCACCCCAAGACGAAGAAAGTCTTGTATATATATGCGGGCGAAGTGGCT	265
Db	1	GTGAAACTCACCCCAAGACGAAGAAAGTCTTGTATATATATGCGGGCGAAGTGGCT	60
Qy	266	AGAAAGCGAAAGCAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT	325
Db	61	AGAAAGCGAAAGCAGAGGGCTTAAAGCTCAATCAACCGAAGCATTGCTTACATTAGT	120
Qy	326	GCCCATATTATGACGAAGCGCGCTGGAAAAAACAACCGTTGCCAGCTTATGGAAGAG	385

1201	DB	CATCACTCCGGGCAAAATGGBAACTTGGCACCGCATGTTGGCGCGAGCTGAAGAGATTTCTAT	1261
1466	QY	GAATGTGGGCTTTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAAAGT	1525
1261	DB	GAATGTGGGCTTTTTTGGGCAAGGCAATAGCTCCAGTAAAAACAACCTCGTAGNACNAGT	1320
1526	QY	AGAAGCGGCGCGAATGGTTTTAAATTGCATGAAGAATCGGGGCACAACAACAAGTGGCAT	1585
1321	DB	AGAAGCGGCGCGAATGGTTTTAAATTGCATGAAGAATCGGGGCACAACAACAAGTGGCAT	1380
1586	QY	CGATCACTGCTTGAGCGCTGGCAGATGAATACGATGTGCAAGTTTGTATTCACACCCGNATC	1645
1381	DB	CGATCACTGCTTGAGCGGTAGCAGATGAATACGATGTGCAAGTTTGTATTCACACCCGNATC	1440
1646	QY	AGTCAATGAGGCAAGTTATGTAGATGACACCTTAAATGCAATGAACGGGCGCGCATCCA	1705
1441	DB	GGTCAATGAGGCAAGTTATGTAGATGACACCTTAAATGCAATGAACGGGCGCGCATCCA	1500
1706	QY	TGCCTTCAACATTTGAGGAGCGGTGGAGACACTCACTGATGTTATATCCATATGCGAGG	1765
1501	DB	TGCCTTCAACATTTGAGGAGCGGTGGAGACACTCACTGATGTTATATCCATATGCGAGG	1560
1766	QY	CGAGCTCAATATTTCTACCCCTCTCCACCAACCCCACTATTTCCCTATACCAATTAACGGT	1825
1561	DB	CGAGCTCAATATTTCTACCCCTCTCCACCAACCCCACTATTTCCCTATACCAATTAACGGT	1620
1826	QY	TGCAGAACACTTAGACATGCTCATGATGCCACCACTAGACAACAGCATCCGCGAGGA	1885
1621	DB	TGCAGAACACTTAGACATGCTCATGATGCCACCACTAGACAACAGCATCCGCGAGGA	1680
1886	QY	TTTACAATTTTCTCAAGCGGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGTCCA	1945
1681	DB	TTCTCCAGTTTTCCAAAGCGGTATCCGCCCGGCTCTATTTGCCGCTGAAGATGTGTCCA	1740
1946	QY	TGATATCGGTGTGATCGCGATGACAAGCTCGAATTCGCAAGCAATGGGGGTCGAGGCGA	2005
1741	DB	TGATATTTGGGCTGATCGCGATGACAAGCTCGAATTCGCAAGCAATGGGGGTCGAGGCGA	1800
2006	QY	AGTGATTCCTCGAACTTTGGCAGACTCGGATAGATAAAGAAATTTGGTAAGCTTCC	2065
1801	DB	AGTGATTCCTAGAACTTTGGCAAACTGGACAAGAATAAAGAAATTTGGTAAGCTTCC	1860
2066	QY	TGAAGATGGCAAAAGATAACGATAAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT	2125
1861	DB	TGAAGATGGTGACAGATAATGACAACTTCCGCATCAACGCTATATCTCCAAATACACTAT	1920
2126	QY	CAACCCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT	2185
1921	DB	TAATATCCCGCTTTGACCCATCGGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT	1980
2186	QY	CGCGCACTTGGTGGTGGAAATCCTCGCTTTTTTGGCGTAAACCCAAATCGTGNATCAA	2245
1981	DB	CGCGCACTTGGTGGTGGAAATCCTCGCTTTTTTGGCGTAAACCCAAATCGTGNATCAA	2040
2246	QY	AGGCGGTATGGTGGTCTTCTCTGAAATGGGCGAATCTTAACGGGCTGTGTGCCCACTCCCA	2305
2041	DB	AGGCGGTATGGTGGTCTTCTCTGAAATGGGCGAATCTTAACGGGCTGTGTGCCCACTCA	2100
2306	QY	ACCGGTTTATTAACCGGAAATGTTTGGGCAATCAACGCAAGGCGAAATTTGAACACGAGCAT	2365
2101	DB	GC CGGTTTATTAACCGGAAATGTTTGGGCAATCAACGCAAGGCGAAATTTGAACACGAGCAT	2160
2366	QY	CACTTTTGTTTCCAAAGTCGCTATGAAAATGGCGTGAAGAAAAGCTGGGCTTAGAGCG	2425
2161	DB	CACTTTTGTTTCCAAAGTCGCTATGAAAATGGCGTGAAGAAAAGCTTAGAGCG	2220
2426	QY	CCAAGTCTTACCGGTCAAAAACTGCCGTAAACTCAACCAAGAAAGCTTCAAGTTCAACGA	2485
2221	DB	CAAGGTGCTACCGGTCAAAAACTGCCGTAAACTCACTAAGAAAGACTTCAAAATTCACAA	2280
2486	QY	CAAAAACGGCAAAATCACCGTCGATCCGAAAACTTTTCGAGGCTTTTGTAGATGGCAAACT	2545
2281	DB	CAAGAAGGCGCATATCACTGTGATCCTTAAACCTTTGAGGCTTTTGTAGATGGCAAACT	2340

Qy 2546 CTGACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGGTACACCTTTCTTCTTAGGC 2605
 |||||
 Db 2341 CTGCACCTCTAAACCCCGCTCTGAAGTGCCTCTAGCCCAACGCTACCTTTCTTCTTAGGC 2400
 |||||
 Qy 2606 ACAAT 2610
 |||||
 Db 2401 ACAAT 2405
 |||||
 RESULT 6
 AX356686
 LOCUS AX356686 2405 bp DNA linear PAT 13-FEB-2002
 DEFINITION Sequence 4 from Patent EP1176192.
 ACCESSION AX356686
 VERSION AX356686.1 GI:18674023
 KEYWORDS
 SOURCE
 ORGANISM
 Helicobacter felis
 Helicobacter felis
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 Helicobacteraceae; Helicobacter.
 1
 Kusters, J.G. and Cattoli, G.
 TITLE Helicobacter felis vaccine
 JOURNAL Patent: EP 1176192-A 4 30-JAN-2002;
 Akzo Nobel N.V. (NL)
 FEATURES
 source
 1..2405
 /organism="Helicobacter felis"
 /mol_type="unassigned DNA"
 /db_xref="taxon:214"
 1..681
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD23257.1"
 /db_xref="GI:18674024"
 /translation="MKLTPEQEKFLYYAGEVARRKKAEGKLKQPEALAIYASIHIM
 DBARRKKTVAIMEECHEHFLKKDVRMGVGNVPDLGVETFPDGTKLTVTMFGGT
 DBHFKAGEYKFGCKDIELNAGKCEVTELENGEKSLHVGSHFHFETNKALKDFDE
 KAYGKRLDIPSGNTLRIAGQTRKYQLIPLGGSKVI GMNGLVNNIADERHKKALKD
 KASHGFLK"
 692..2398
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD23258.1"
 /db_xref="GI:18674025"
 /translation="MKMKQEVVYTGTTGDKVRLGDTDLWAEVEHDYTTYGEELKPF
 GAGKTIREGMGQSNPDENLDLVITNAMI IDYTG IYKADIGIKNGKHIGIKGAKND
 MDQGVPMHVVGVGTWALAGEMII TAGGIDSHTFHFLSPQQPPTALANGVTVMFGGT
 GPGVDGNATITPFGNHLHRLMRAAEYSMNVFLGKGNSSKKQLQVEVGAAGICFK
 LHEDGTTPTADICHLNSVADEVQVCIHTDITNEAGYVDITLNMNGRAIHAYHIEG
 AGGSHSPDITWAGLNIIPSTTPTI PVTINTVAEHLDMLTCHLDKRIREDIQLPS
 QSRIRPGSTAAEDVLHDIGVIAMTSSDSQAMGRAGEVIPRTWQTADKNKKKFGKLPEP
 GADNDFRIKRYISNTIPNPALGVSEYIGVSEEGKIADLVVWPAFGVAKPKIVIK
 GGMVVFSEMGDSVPTPPQVYYREMFHGHHGAKFDTSITFVSKPVAEYGVKELGL
 ERKVLVPVNCRNITKKDFKFNKTAHTITVDPKTFEYVDGKLCITSKPASAEVLPQRYT
 PF"

ORIGIN

```
Query Match      74.9%; Score 2158.6; DB 2; Length 2405;
Best Local Similarity 93.6%;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
```

206	GTGAAACTCACCCAAAGAGCAAGAAAGTCTTGTTATATATATGCGGCGCAAGTGCT	265
Qy		
1	GTGAAACTCACCCAAAGAGCAAGAAAGTCTTGTTATATATGCGGCGCAAGTGCT	60
Db		
266	AGAAAGCGCAACGACAGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTTACATTAGT	325
Qy		
61	AGAAAGCGCAACGACAGAGGGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTTACATTAGT	120
Db		

QY 326 GCCCATATTATGAGAGAGCGCGCGTGGAAAAAACCCTGTCGCCAGCTTATGGAAGAG 385
DB 121 GCCCATATTATGAGAGAGCGCGCGTGGAAAAAACCCTGTCGAACTTATGGAAGAA 180
QY 386 TGCATGCACTTTTGA AAAAGATGAAGTAATGCCGGGTGGGTAAATATGTTCCCGAT 445
DB 181 TGTATGCATTTTGA AAAAGATGAAGTAATGCCGGGTGGGTAAATATGTTCCCGAT 240
QY 446 CTAGGTGTAAGGCCACTTTCTGATGGTACGAACTTTGAATCTGTAATTTGGGCCATC 505
DB 241 TTGGGCGTGAAGCCACTTTCCCGATGGCACCAACTCGTAAACCGTGAATTTGGCCAT 300
QY 506 GAACCAAGTAGCACTTCAAGCGGCGGAAGTGAATTTGGTGGGATAAAGACATCGAG 565
DB 301 GAACCTGATGAACACTTTAAAGCGCGTGAAGTGAATTTGGTGGGATAAAGACATTCGAG 360
QY 566 CTCATGCGAGCAAGAGTAACCGACTTCAGGTTTACTAATGAAGGCCCTAAATCTCTTG 625
DB 361 CTCACGCGGTGAAGAGTTACCGAGCTTGAAGTTTACAAGAGGACCTTAAATCTCTTG 420
QY 626 CATGTGGGTAGCCATTTCCACTTTCTTGAAGCTAAACAGGCACTAAAAATTCGATCGTGA 685
DB 421 CATGTGGGTAGCCATTTCCACTTTCTTGAAGCAACAGGCACTTGAATTCGATCGGGA 480
QY 686 AAAGCTTANGCAAAACGCTTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA 745
DB 481 AAAGCTTANGCAAAACGCTTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA 540
QY 746 CAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAGTGAATTTGGCATGAAC 805
DB 541 CAACCCGCAAGTGCAGTTGATTCCTCTTGGCGGTAGTAAAAAGTGAATTTGGCATGAAC 600
QY 806 GGGCTGTGAAATAACATCGCGGATGAACGCGCATAAACATAAAGCGCTTGAACAGGCGAAA 865
DB 601 GGGCTGTGAAATAATTTGCGGACGAAGCCCAATAACACAAAGCACTAGACAGGCAAAA 660
QY 866 TCTCAGCGATTTATCAAGTAAGAGAGATCCCATGAAATGAAAAACAAAGAAATATGTAAA 925
DB 661 TCTCAGCGATTCATCAAGTAAGAGAGATCCCATGAAATGAAAAACAAAGAGTATGTAAA 720
QY 926 TACTACGGACCCACCAAGGCGATAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGA 985
DB 721 CACCTACGGACCCACCAAGGCGATAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGA 780
QY 986 AGTAGAACATCACTATACCATTTAGGCGAAGAACTTAAATTTGGCGGGTAAAAACTAT 1045
DB 781 AGTAGAACATCACTATACCATTTAGGCGAAGAGCTCAAAATTTGGCGGGTAAAAACTAT 840
QY 1046 CCGTGAGGATATGGGTACAGACCAATAGCCAGATGAAAAACACCTTAGTATTTAGTCATCAC 1105
DB 841 CCGTGAGGATATGGGTACAGACCAATAGCCAGATGAAAAACACCTTAGTATTTAGTCATCAC 900
QY 1106 TAACCGGATGATATGCACTACACCGGATTTAACAAGCGGACATTTGGGATTTAAAAACGG 1165
DB 901 CAACCGGATGATATGCACTACACCGGATTTTATAAGCGGACATTTGGTATTTAAAAATGG 960
QY 1166 CAATAATCATCGCATTTGCAAGGAGGAAACAGGACATGCAAGATGCGTAAGCCCTCA 1225
DB 961 CAATAATCATCGCATTTGCAAGGAGGAGAAACAGGACATGCAAGATGCGTAAGCCCTCA 1020
QY 1226 TATGTCGTGGGTGTCAGACAGCACTAGCAGGGAAGGTATGATTTATACCGCTGG 1285
DB 1021 TATGTCGTGGGTGTCAGACAGCACTAGCAGGGAAGGTATGATTTATACCGCTGG 1080
QY 1286 GGGAAATCGATTTACACACCCACTTCTCTTTCTCCACAACAATTTCCCTACCGCTCTAGCCAA 1345
DB 1081 GGGGATCGATTTCCGACACCCACTTCTCTCTCTCCCAACAAATTTCCCTACCGCTCTAGCCAA 1140
QY 1346 TGGCGTTACCAACCATTTTGGAGGCGGCACAGGCTCTGTAGATGCGAATGCGACTAC 1405
DB 1141 TGGGTGTTACCAACCATTTTGGAGGTGGCACAGGTCGGGTAGATGCAAGATTCGACAC 1200
QY 1406 TATCACTCCGGGCAAAATGGAACCTTGACCGCATGTTGCGCGCAGCAGAGAGTATTTCTAT 1465

DB 1201 CATCACTCCGGGCAAAATGGAACCTTGACCGCATGTTGCGCGCAGCTGGAAGAGTATTTCTAT 1260
QY 1466 GAAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAAACAAGT 1525
DB 1261 GAAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGTAAAAAACAATCTGTAGAAACAAGT 1320
QY 1526 AGAAGCGGCGCGATTTGGTTTAAATTTGCATGAAGACTGGGGCAACCAACCAAGTGGCAT 1585
DB 1321 AGAAGCGGCGCGATTTGGCTTTAAATTTGCATGAAGACTGGGGCAACCAACCAAGTGGCAT 1380
QY 1586 CGATCACTGCTTTGAGCGTGGCAGATGAATACCATGTGCAAGTTTGTATCCACACCGATAC 1645
DB 1381 CGATCACTGCTTTGAGCGTGGCAGATGAATACCATGTGCAAGTTTGTATCCACACCGATAC 1440
QY 1646 AGTCAATGAGGCGAGTTATGTAGATGAACCCCTAAATGCAATGAACGGGCGGCCATCCA 1705
DB 1441 GGTCAATGAGGCGAGTTATGTAGATGAACCCCTAAATGCGATGAACGGGCGGCCATCCA 1500
QY 1706 TGCCTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATGCGAGG 1765
DB 1501 TGCCTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATGCGAGG 1560
QY 1766 CGAGCTCAATTTCTACCTCTCTCACCAACCCCGCACTATTCCTTATACCATTAATACGGT 1825
DB 1561 CGAGCTCAATTTCTACCTCTCTCACCAACCCCGCACTATTCCTTATACCATTAATACGGT 1620
QY 1826 TGCAGAACCACTTTAGACATGCTCATGACATGCCACACCTAGACAAACGATCCGCGAGGA 1885
DB 1621 TGCAGAACCACTTTAGACATGCTCATGACCTGCCACCACTAGACAAACGATCCGCGAGGA 1680
QY 1886 TTTCACAAATTTCTCAAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGTCCA 1945
DB 1681 TCTCCAGTTTTCCTCAAAAGCGGTATCGCCCGGCTCTATTCGCGCTGAAGATGTGTCCA 1740
QY 1946 TGATATGGTGTGATCGCGATGCAAGCTCGGATTTGCGAAGCAATGCGGGCTGCGAGCGA 2005
DB 1741 TGATATTTGGCGTGTGCGGATGCAAGCTCGGATTTGCGAAGCAATGCGGGCTGCGGGA 1800
QY 2006 AGTGAATCTCTCGAACTTGCACACTGCGGATGAAGATAAAGAAATTTGGTAAAGCTTCC 2065
DB 1801 AGTGAATCTCTAGAACTTGCACAACTTGCAAGAAATAAAGAAATTTGGTAAAGCTTCC 1860
QY 2066 TGAAGATGCGCAAGATAACGATAATTTTCGCAATTAAGCGCTACATCTCCAAATACACTAT 2125
DB 1861 TGAAGATGCTGAGATAATGACAACTTTCGCACTCAAACTATATCTCCAAATACACTAT 1920
QY 2126 CAAACCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT 2185
DB 1921 TAATCCCGCTTTGACCCCATGGCGTGAAGCGATATATCGGCTCTGTGGAAGAGGCGCAAGAT 1980
QY 2186 CGCCGACTTGTGTGTGGAATCTCGCTTTTGTGGCTGAAGAAACCCAAATCTGTATCAA 2245
DB 1981 CGCCGACTTGTGTGTGGAATCTCGCTTTTGTGGCTGAAGAAACCCAAATCTGTATCAA 2040
QY 2246 AGCGGCTATGTTGTGCTTCTCTGAAATGGGCGATTTCTAAACCGGCTCTGTGCCACTCCCA 2305
DB 2041 AGCGGCTATGTTGTGCTTCTCTGAAATGGGCGATTTCTAAACCGGCTCTGTGCCACTCCCA 2100
QY 2306 ACCGGTTATTTACCGCGAAATTTTGGGCAATCAAGGCAAGGCGAAATTTTGACACCGAGCAT 2365
DB 2101 GCGCGGTTTATTAACCGCGAAATTTTGGGCAATCAAGGCAAGGCGAAATTTTGACACCGAGCAT 2160
QY 2366 CACTTTTGTTCGAAGTGGCTATGAAATGGGCTGAAGAAAGCTGGCTTTAGAGCG 2425
DB 2161 CACTTTTGTTCGAAGTGGCTATGAAATGGGCTGAAGAAAGCTGGCTTTAGAGCG 2220
QY 2426 CCAAGTCTTACCGGTCAAAAACCTGCGTAAACATCACAGAAAGAGCTTCAAGTTTCAACGA 2485
DB 2221 CAAAGTGTCTACCGGTGAAAAACCTGCGCAACATCATAGAAAGAGCTTCAAAATTCACAA 2280
QY 2486 CAAAAACGCGCAAAATCAACCGTGCATCCGAAAAACCTTTCGAGGCTTTTGTAGATGGCAAACT 2545

```

Db      2281 CAAGACGGCGCATATCACTGTGATCTATAAACCTTCGAGGTCTTTGTAGATGGCAACT 2340
Qy      2546 CTGCACCTCTAAACCCACCTCGAAGTGCCTCTAGCCAGCGCTACACTTCTTCTAGGC 2605
Db      2341 CTGCACCTCTAAACCCCGCTCTGAAGTGCCTCTAGCCAGCGCTACACTTCTTCTAGGC 2400
Qy      2606 ACAAT 2610
Db      2401 ACAAT 2405

RESULT 7
BD185305
LOCUS   BD185305                2407 bp    DNA        linear    PAT 17-JUN-2003
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185305
VERSION   BD185305.1 GI:31877505
KEYWORDS JP 2002355054-A/4.
SOURCE   Helicobacter felis
ORGANISM Helicobacter felis
          Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
          Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2407)
AUTHORS   Kusters,J.G. and Cattoli,G.
TITLE     Helicobacter felis vaccine
JOURNAL   Patent: JP 2002355054-A 4 10-DEC-2002;
          AKZO NOBEL NV
COMMENT   OS Helicobacter felis
          PN JP 2002355054-A/4
          PD 10-DEC-2002
          PF 16-JUL-2001 JP 2001214711
          PR 17-JUL-2000 EP 00202565.8
          PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI
          PC
          C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
          A61K39/23,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
          PC C12N1/19,
          C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC
          53,
          PC
          G01N33/53,G01N33/566,G01N33/569//(C12N9/80,C12R1/01),(C12Q1/68, PC
          C12R1/01),
          PC C12N15/00,C12N5/00,A61K37/02
          CC Helicobacter felis vaccine
          FH Key Location/Qualifiers
          FT CDS (2)..(682)
          FT CDS (693)..(2399).
FEATURES
    source          1..2407
                     /organism="Helicobacter felis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:214"

ORIGIN
Query Match 74.9%; Score 2158.6; DB 2; Length 2407;
Best Local Similarity 93.6%; Pred.No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy      206 GTGAACCTCACACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGGCT 265
Db      2 GTGAACCTCACACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGGCT 61
Qy      266 AGAAGCGCAAGACGAGGCTTAAAGCTCAACCAACCCGAGCCATGTGCTTACATTAAGT 325
Db      62 AGAAGCGCAAGACGAGGCTTAAAGCTCAACCAACCCGAGCCATGTGCTTACATTAAGT 121
Qy      326 GCCCATATATGACGACGCGCGTGGAAAAAACAACCGTCCAGCTATGTTGAAGAG 385
Db      122 GCCCATATATGACGAGGCGCGCGTGGAAAAAAGACCGTTCGGAACTTATGGAAGAG 181
Qy      386 TGCAATGCACTTTTGTAAAAAAGATGAAGTAATGCCGGGGTGGTAAATATATGTTCCCGAT 445

```

```

Db      182 TGTATGCACTTTTGAANAAGACGAGGTGATGCCCGGTGTGGGGAATATGGTCCCTGAT 241
Qy      446 CTAGGTAGTAAGACCACTTTCTCTGATGTTAGCAAACTTGTAACTGTGAATTTGGCCCATC 505
Db      242 TTAGGCGTGAAGCTACTTTTCCCGATGACCAAACTCGTAACCGTGAATTTGGCCCATC 301
Qy      506 GAAACAGATGACACTTTCAAAGGGGCGAAGTGAATTTGGTTCGATATAAAGACATCGAG 565
Db      302 GAAACCGATGAACACTTTCAAAGCGGCGAAGTCAAAATTTGGCTGTGTATAAAGACATTTGAA 361
Qy      566 CTCATGCGAGCAAGAAGACTTAACCACTTGAAGTTACTAAATGAAGGCGCTAAATCCTTG 625
Db      362 CTCACGCGAGTGAAGAGTTTACCGAACTAGAAGTTACCACGAAGGACCTAAATCCTTG 421
Qy      626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACCAAGGCCTAAATTTGATCGTGA 685
Db      422 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCAACAAGGCATTGAATTCGATCGGAA 481
Qy      686 AAGCCTATGGAACCGCTAGATATTCCTCTGGCAACAACGCTACGATTTGGGGCAGGA 745
Db      482 AAGCCTATGGAACCGCTAGATATTCCTCTGGCAACAACGCTACGATTTGGGGCAGGA 541
Qy      746 CAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATTCGATGAAC 805
Db      542 CAACCCGTAAGTGCAGTTAATTCCTCTTGGCGGCGAGTAAAAAGTGAATTCGATGAAC 601
Qy      806 GGGCTTGTGAATAACATCGCGATGAACCGCCATAACCAATAAAGCGCTTCAAGAGGCGAAA 865
Db      602 GGGCTTGTGAATAATATTGCGAGTGAACCGCAATAACCAAGCGTTAGAAAAGCAAAA 661
Qy      866 TCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAAAACAAGAAATATGTAAA 925
Db      662 TCTCAGGATTTATCAAAATAGGAGACTCCCATGAAATGAAAAACAAGAGATATGTAAA 721
Qy      926 TACCTACGCAACCCCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGA 985
Db      722 TACCTACGCAACCTTACCACAGGCGCAAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGA 781
Qy      986 AGTAGAATGACTATATACCATATGCGCAAGAACTTAAATTTGGCGGGTAAAAACTAT 1045
Db      782 AGTAGAATGACTATATACCATATGCGCAAGAGCTTAAATTTGGCGGGTAAAAACTAT 841
Qy      1046 CCGTGAGGCTATGGTTCAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCAATCAC 1105
Db      842 CCGTGAGGCGATGGTTCAGAGCAATAGTCCAGATGAAAAACCCCTAGATTTAGTCAATCAC 901
Qy      1106 TAAACGCGATGATTATCGACTACACCGGATTTTAAAGCCGACATTTGGGATTTAAACCGG 1165
Db      902 CAACGCGATGATTATGACTACACCGGATTTTAAAGCCGACATTTGGGCAATTTAAAAATGG 961
Qy      1166 CAAATCCATGGCATTTGCAAGGCGAAGAAACAAGCATGCAAGATGGCGTGAAGCCCTCA 1225
Db      962 CAAATCCATGGCATTTGCAAGGCGAAGAAACAAGCATGCAAGATGGCGTGAAGCCCTCA 1021
Qy      1226 TATGTCGTGGGTGTGGGCACAGAAGCACTAGCAGGGAAGGTATGATTTATACCGCTGG 1285
Db      1022 TATGTCGTGGGTGTGGGCACAGAAGCATTAGCAGGGAAGGTATGATTTATACCGCTGG 1081
Qy      1286 GGGAAATCGATTACACACCCACTTCTCTTTCACAAATTTCCCTACCGCTCTAGCCAA 1345
Db      1082 GGGGATCGATTACACACCCACTTCTCTCTCCACAACTTTCCCTACCGCTCTAGCCAA 1141
Qy      1346 TGGCGTTACACCATGTTTGGAGGCGGCACAGGTCTCTAGATGCGCAAGATCGCACTAC 1405
Db      1142 TGGCGTTACACCATGTTTGGCGGTGGGCACAGGTCCGGTAGATGCGCAAGATCGCACTAC 1201
Qy      1406 TATCACTCCGCGCAAAATGGAACCTTGACCGCATTTGCGCGCAGCAGAGAGTATTCTAT 1465
Db      1202 CATCACTCCGCGCAAAATGGAACCTTGCAACCGCATTTTGGCGCAGCTGAAGAGTATTCTAT 1261
Qy      1466 GAATGTGGGCTTTTGGGCAAGAGCAATAGCTCTTAGCAAAAAACAACCTTGTAGAACCAAGT 1525

```

1262	GAATGTGGGCTTTTGGGCAAGCGCAATAGCTCCAGTAAAAAACAACCTGTAGAACAAAT	1321
1526	AGAAGCGGGCGCGATTTGGTTTTTAAATTTGCATGAAGACTGGGGCACAAACCAAGTGCAGT	1585
1322	AGAAGCGGGCGCGATCGGCTTTTAAATTTGCATGAAGACTGGGGCACAACTCCAAGTGCAT	1381
1586	CGATCACTGCTTGAAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1645
1382	CGATCACTGCTTGAAGCGTAGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1441
1646	AGTCAATGAGGCGAGTTATGTAGATGACACCCCTAAATGCAATGAAACGGGCGGCCATCCA	1705
1442	GGTCAATGAGGCGAGTTATGTAGATGACACCCCTGAATGCGCATGAAACGGGCGGCCATCCA	1501
1706	TGCTTACCACATTTGAGGAGCGGGTGAGGACACTCACCTGATGTTATCACCATGCGAGG	1765
1502	TGCTTACCACATTTGAGGAGCGGGGAGGAGCACTCACCTGATGTTATCACCATGCGAGG	1561
1766	CGAGCTCAATTTCTACCTCTTCCACCAACCCCACTATTTCCCTATACCAATTAATACGGT	1825
1562	CGAGCTCAATTTCTACCTCTTCCACCAACCCCACTATTTCCCTATACCAATTAATACGGT	1621
1826	TGCAGAACCTTAGACATGCTCATGATCGCACCCACCTAGACAAACGGCATCCGCGAGGA	1885
1622	TGCAGAACCTTAGACATGCTCATGATCGCACCCACCTAGATTAACGGCATCCGCGAGGA	1681
1886	TTTACAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGTCCA	1945
1682	TTTACAATTTTCCAAAGCCGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGTCCA	1741
1946	TGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGTGACGGCGA	2005
1742	TGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGTGACGGCGA	1801
2006	AGTGATTCCTCGAACTTGGCAGACTCGGATAGAAATAAAAAGAAATTTGGTAAAGCTTCC	2065
1802	AGTGATTCCTCGAACTTGGCAGACTCGGATAGAAATAAAAAGAAATTTGGTAAAGCTTCC	1861
2066	TGAAGTGGCAAGATAACGATAAATTTCCGATTTAAGCGCTACATCTCCAAATACACTAT	2125
1862	TGAAGATAGTGAGATTAACGAACTTCCGATTCGAACGGCTACATCTCCAAATACACTAT	1921
2126	CAACCCCGCTTTGACCCACCGCGTGACGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT	2185
1922	TAACCCCGCTCTAACCCATCGGGTAAAGAGTATATCGGCTCTGTGGAAGAGGGCAAAAT	1981
2186	CGCGACTTGGTGTGGAATCCTGCTTTTTTTGGCGTAAACCCAAAATTCGTATCAAA	2245
1982	CGCTGATTTGGTGTGGAATCCTGCTTTTTTTGGTGTGAAACCTAAGATTTGTATCAA	2041
2246	AGGCGGTATGTTGTTCTCTGAAATGGGCGATTTCTAACCGCTCTGTGCCACTCCCCA	2305
2042	AGGCGGTATGTTGTTCTCTGAAATGGGCGACTCCAACCGCTCGGCTACACCTCA	2101
2306	ACCGGTTTATTACCGCGAAATGTTTGGGCGATCACGGCAAGCGGAAATTTGACACCGCAT	2365
2102	GCCGGTTTATTACCGCGAAATGTTTGGGCGATCACGGCAAGCGGAAATTTGACACCGCAT	2161
2366	CACTTTGTTCAAAGTCGCTATGAAATTTGGCGTGAAGAAAGCTGGGCTTAGAGCG	2425
2162	CACTTTGTTCAAAGTCGCTATGAAATTTGGCGTGAAGAAAGAACTAGGCTTAGAGCG	2221
2426	CCAGTTTCTACCGGTCAAAAACCTGCCGTAACTCACCAAGAAACATTTCAAGTTCAACGA	2485
2222	CAAGGTGCTACCGGTGAAAAACCTGCCGCAACATCACTAAGAAAGACTTCAAAATCAACAA	2281
2486	CAAAAACGGCAAAAATCACCGTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGGCAACT	2545
2282	CAAGACGGCGCATATCACTGTGATCCTTAAACCTTCGAGGTCTTTGTAGATGGCAACT	2341
2546	CTGACCTCTAAACCCCACTCGCAAGTGCCCTTAGAGCCGCGTACACTTTCTTTAGGC	2605
2342	CTGACCTCTAAACCCGCTCTGAAGTGCCCTTAGCGCAGCGCTACACTTTCTTTAGGC	2401

[illegible]

386 TGCATGCACCTTTTGCARAAAGATGACGTAATGCCCAGGCTGGGTAAATATATGGTTCCCGAT 445
182 TGTATGCACCTTTTGCARAAAGAGAGAGGTGATGCCGGTGGGGAAATATGGTCCCTGAT 241
446 CTAGGTGTAGAAAGCACCTTTCTGATGGTACGAAACTTGTAACTGTGAAATTTGGGCCCATC 505
242 TTAGCGGTGGAAGTACTTTTCCCGATGGCACAAACTCGTAACCGTGAATTTGGCCCATC 301
506 GAACACAGATGAGCACTTCAAAAGCGGGCGAAGTGAATTTGGTTCGGATAAAGACATCGAG 565
302 GAACCCGATGAACACTTCAAAAGCGGGCGAAGTCAAAATTTGGCTGTGATAAAGACATTCGAA 361
566 CTCAATGACGCAAGAAGTAAACCGAATCTGAGTTACTAATGAAGGCCCTAAATCTCTTG 625
362 CTCAACCGAGTAAGGAAGTTACCGAATAGAGTTTACCAACGAAGGACCTTAAATCTCTTG 421
626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAAACGAAGGCACATAAAATTCGATCGTGAA 685
422 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCCAAACGAAGCATTTGAAATTCGATCGGGAA 481
686 AAAGCCTATGSCAAACGCTAGATATTTCCCTCTGGCACAACGCTACGCAATTTGGGCGAGGA 745
482 AAAGCCTATGSCAAACGCTAGATATTTCCCTCTGGCACAACGCTACGCAATTTGGGCGAGGA 541
746 CAACCCGCAAGTGCAGTTGATTCCTCTGGTGGCAGTAAAGATGGAATTTGGGCGATGAAC 805
542 CAACCCGCAAGTGCAGTTAATCCCTCTGGGCGCAGTAAAGATGGAATTTGGGCGATGAAC 601
806 GGGCTTGTGAATACATCGCGGATGAACGCGCATAAACATAAAGCGCTTGAACAGGCGGAA 865
602 GGGCTTGTGAATATTTGCAAGTGAACGCGCATAAACATAAAGCGCTTGAACAGGCGGAA 661
866 TCTCAGCGATTTATCAAGTAAAGAGACTCCCATGAAATGAAAAACAAAGATATGTAAA 925
562 TCTCAGCGATTTATCAATTAAGAGACTCCCATGAAATGAAAAACAAAGATATGTAAA 721
926 TACCTACGACCCACCAAGCGGATAAAGTCGCTTTAGAGATACCGATCTTTGGGCGAGA 985
722 TACCTACGACCCACCAAGCGGACAAAGTCGCTTTAGGAGATACCGATCTTTGGGCGAGA 781
986 AGTAGACATGACTATACCACTATGCGGAAGAACTTTAAATTTGGCGGGGTAAACCTAT 1045
782 AGTAGAACATGACTATACCACTTATGCGGAAGAGCTCAAAATTTGGCGGGGTAAACCTAT 841
1046 CCGTAGGGTATGGGTACAGACCAATAGCCCTGTATGAAAAACACCCCTAGATTTAGTCATCAC 1105
842 CCGTAGGGCATGGGTACAGACCAATAGTCAGATGAAAAACACCCCTAGATTTAGTCATCAC 901
1106 TAAACGCGATGATTATCGACTACACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG 1165
902 CAACGCGATGATTATTTGACTATACACCGGGATTTACAAAGCCGACATTTGGCATTAAAAATGG 961
1166 CAATAATCATGGCATTTGGCAAGGCGGAAACAGGGAATGCAAGATGGGTAAAGCCCTCA 1225
962 CAATAATCATGGCATTTGGCAAGGCGGAAACAGGGAATGCAAGATGGGTAAAGCCCTCA 1021
1226 TATGTGTGGGTGGGTCAGAGACATAGCAGGGGAGGTATGATTTATCCGCTGG 1285
1022 TATGGTGTGGGTGGGTCAGAGACATTTAGCAGGGGAGGTATGATTTATCCGCTGG 1081
1286 GGGAAATCGATTATCACACCCACTTCTTTCTTCCACAAATTTCCCTACCGCTCTAGGCCAA 1345
1082 GGGGATCGATTATCACACCCACTTCTTCTTCCACAAATTTCCCTACCGCTCTAGGCCAA 1141
1346 TGGGTTTAAACCATGTTTGGAGCGGCAACAGTCTCTGTAGATGGCAAGATTCGACTAC 1405
1142 TGGCGTTTAAACCATGTTTGGCGGTGGCAACAGTCTCCGTTAGATGGCAAGATTCGACTAC 1201
1406 TATCACTCCGGCAATGGAATCTTGACCGCATTTGGCGGCGACAGAGATTTCTAT 1465
1202 CATCACTCCGGCAATGGAATCTTGACCGCATTTGGCGGCGACAGAGATTTCTAT 1261
1466 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAAACAAGT 1525

1262 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCCAGTAAAAACAACCTTGTAGAACAAAT 1321
1526 AGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAGACTGGGGCACACCAACAGTTCGAT 1585
1322 AGAAGCGGGCGGATCGGCTTTAAATTTGCATGAAGACTGGGGCACAACTCCAAGTGCAT 1381
1586 CGATCACTGCTTTAGCGTGGCAGATGAATACGATGTGGAAGTGTGTATCCACACCGATAC 1645
1382 CGATCACTGCTTTAGCGTGGCAGATGAATACGATGTGGAAGTGTGTATCCACACCGATAC 1441
1646 AGTCAATGAGGCGAGTTATGTAGATGACACCTTAAATGCAATGAAACGGGCGGCGCATCCA 1705
1442 GGTCAATGAGGCGAGTTATGTAGATGACACCTTGAATGCGATGAACGGGCGGCGCATCCA 1501
1706 TGCTTACCACATTTGAGGAGCGGTTGAGGAGCACTCACCTGATGTTATCACCATTTGGCAGG 1765
1502 TGCTTACCACATTTGAGGAGCGGTTGAGGAGCACTCACCTGATGTTATCACCATTTGGCAGG 1561
1766 CGAGCTCAATATTTCTACCTCTCCACACCCCTATATCCCTATATCCATTAATACGGT 1825
1562 TGACAGAACACTTTAGACATGCTCATGACATGCCACCACTTAGACAAACGCAATCCGCGAGGA 1885
1622 TGACAGAACACTTTAGACATGCTCATGACCTGCGCACCACTTAGATAAACGCAATCCGCGAGGA 1681
1886 TTTTCAATTTTCTCAAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCA 1945
1682 TTTTCAATTTTCTCAAAAGCGGTATCGCCCGGCTCTATCGCCCGGCTGAAGATGTGCTCCA 1741
1946 TGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGGA 2005
1742 TGATATTTGGCGTGTATCGCGATGCAAGCTCGGATTCGCAAGCAATGGGCGGCTGCAGGA 1801
2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTGAAGCTTCC 2065
1802 AGTGATTTCTCGAACTTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTGAAGCTTCC 1861
2066 TGAAGATGCGAAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATPACACTAT 2125
1862 TGAAGATGCGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATPACACTAT 1921
2126 CAACCCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT 2185
1922 TAAACCCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAAAT 1981
2186 CGCGCACTTTGGTGTGGAATCTTGGCTTTTGGCGGTAAAAACCCAAATCGTGATCAA 2245
1982 CGCTGATTTGGTGTGGAATCTTGGCTTTTGGCGGTAAAAACCCAAATCGTGATCAA 2041
2246 AGGCGGTATGGGTCTTCTCTGAAATGGGCGATTTCAACGCGTCTGTGCCCACTCCCA 2305
2042 AGGCGGTATGGGTCTTCTCTGAAATGGGCGACTTCAACGCGTCTGTGCCCACTCA 2101
2306 ACCGGTTATTTACCCCGGAAATGTTTGGGCGCATCAACGCGAGGCGAAATTTGACACCAAGCAT 2365
2102 GCGGTTTATTTACCCCGGAAATGTTTGGGCGCATCAACGCGAGGCGAAATTTGACACCAAGCAT 2161
2366 CACTTTTGTTCCTTCCAAAGTCCGCTATGAAAAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCG 2425
2162 CACTTTTGTTCCTTCCAAAGTCCGCTATGAAAAATGGCGTGAAGAAAAAGCTAGGCTTAGAGCG 2221
2426 CCAAGTTCTACGGGTCAAAAACCTGCGGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGA 2485
2222 CAAGGTGCTACCGGTGAAAAACCTGCGCAACATCACTAAGAAAGACTTCAAAATTTCAACAA 2281
2486 CAAAACGCGCAAAAATCACCGCTGATCCGAAACCTTTCGAGGCTTTTGTAGATGGCAAACT 2545
2282 CAAGACGGCGCATATCACTGTGATCTTAAACCTTTCGAGGCTTTTGTAGATGGCAAACT 2341
2546 CTGACCTCTTAAACCCACCTCGCAAGTGCCTCTAGCCCGAGCGCTACACTTCTTCTTAGGC 2605

Db 2342 CTGCACCTCTAAACCCGGCTCTGAAGTGCCTCTAGCCAGCGCTACACTTCTCTTAGGC 2401

Qy 2606 ACAATG 2611

Db 2402 NCAATG 2407

RESULT 9

BD185304

LOCUS

DEFINITION Helicobacter felis vaccine.

ACCESSION BD185304

VERSION BD185304.1 GI:31877504

KEYWORDS JP 2002355054-A/3.

SOURCE Helicobacter felis

ORGANISM Helicobacter felis

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.

1 (bases 1 to 2183)

AUTHORS Kusters, J.G. and Cattoli, G.

TITLE Helicobacter felis vaccine

JOURNAL Patent: JP 2002355054-A 3 10-DEC-2002;

AKZO NOBEL NV

COMMENT OS Helicobacter felis

PN JP 2002355054-A/3

PD 10-DEC-2002

PF 16-JUL-2001 JP 2001214711

PR 17-JUL-2000 EP 0202565.8

PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI

PC

C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC

A61K39/23,

PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,

PC C12N1/19,

PC

C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC

53,

PC

G01N33/53, G01N33/566, G01N33/569// (C12N9/80, C12R1:01), (C12Q1/68, PC

C12R1:01),

PC C12N15/00, C12N5/00, A61K37/02

CC Helicobacter felis vaccine

FH Key Location/Qualifiers

FT CDS (3)..(683)

FT CDS (694)..(2181).

FEATURES

source

1..2183

/organism="Helicobacter felis"

/mol_type="genomic DNA"

/db_xref="taxon:214"

ORIGIN

Query Match 67.6%; Score 1948; DB 2; Length 2183;

Best Local Similarity 93.3%; Pred. No. 0;

Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 206 GTGAACCTCACACCCAGAGCAAGAAAGTTCTTGTTATATTATGCGGCGAAGTGGCT 265

Db 3 GTGAACCTCACACCCAGAGCAAGAAAGTTCTTGTTATATTATGCGGCGAAGTGGCT 62

Qy 266 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTTACATTAGT 325

Db 63 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCCGAGCCATTGCGCTACATTAGT 122

Qy 326 GCCCATATTATGAGCAAGCGCGCTGGAAAAAACCCTTGGCCAGCTTATGGAAGAG 385

Db 123 GCCCATATTATGAGCAGCGCGCTGGAAAAAACCCTTGGCTCAACTTATGGAAGAA 182

Qy 386 TGCATGCACTTTTCAAAAAAGATCAAGTAAATGCCGGGGTGGTAAATATGTTCCCGAT 445

Db 183 TGTATGCACTTTTCAAAAAAGATGAGGTGATGCCCGGTGGGGAATATGTTCCCTGAT 242

Qy 446 CTAGGTGTAGAGCCACCTTTCTCTGATGCTACGAACTGTGAATGTGAATTTGGCCCATC 505

Db 243 TTGGCGGTAGAGCCACTTTCCCGATGGCACCAAACTCGTAACCGTGAATTTGGCCCAT 302

Qy 506 GAAACAGATGAGCACTTCAAAGCGGGCGAAGTAAATTTGGTTGCATATAAGACATCGAG 565

Db 303 GAACTCTGATGAACACTTTAAAGCCGGTGAAGTAAATTTGGCTGTGATAAAGACATTGAG 362

Qy 566 CTCATATGAGGCAAAAGAGTAAACGAACTTTGAGGTACTTAATGAAGGGGCTAAATCCTTG 625

Db 363 CTCACGTGGGTGAAGAAAGTTACCGAGCTTGAAGTTACCAACGAAGCACTAAATCCTTG 422

Qy 626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCACTAAATTCGATCGTGA 685

Db 423 CATGTGGGTAGCCATTTCCACTTCTTTGAACCAACAGGCACTTGAATTCGATCGGGAA 482

Qy 686 AAAGCTATGGCAACCGCTAGATATTCCTCTGGCAACACGCTACGCACTTGGGGCAGGA 745

Db 483 AAAGCTATGGCAACCGCTAGATATTTCCCTCTGGCAACACGCTACGCACTTGGGGCAGGA 542

Qy 746 CAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATTCGATGAAC 805

Db 543 CAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAGTGAATTCGATGAAC 602

Qy 806 GGGCTTGTGAATAACATCGCGATGAACGCGATAACATAAAGCGCTTGACAGGCGAAA 865

Db 603 GGGCTTGTGAATAATATTTCGGGACGAACGCGCATNAACACAAAGCACTAGACAAGGCAAAA 662

Qy 866 TCTCAGCGATTTATCAAGTAAAGGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAA 925

Db 663 TCTCAGCGATTTATCAAGTAAAGGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAA 722

Qy 926 TACCTACGAGCCACCACAAAGGCGATAAGTGCCTTAGAGATACCGATCTTTTGGGCGAGA 985

Db 723 CACCTACGAGCCACCACCAAGGCGATAAGTGCCTTAGAGATACCGATCTTTTGGGCGAGA 782

Qy 986 AGTAGAACAATGACTATACCACTTATGGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT 1045

Db 783 AGTAGAACAATGACTATACCACTTATGGGCGAAGAGCTCAAATTTGGCGCGGTAAAACTAT 842

Qy 1046 CCGTCAGGGTATGGGTGAGAGCAATAGCCCTCATGAAAAACCCCTAGATTAGTCAATCAC 1105

Db 843 CCGTCAGGGTATGGGTGAGAGCAATAGCCCTCATGAAAAACCCCTAGATTAGTCAATCAC 902

Qy 1106 TAAACGCGATGATTATCGACTACACCGGGATTTACAAAGCCGACATTTGGGATTAACAAACGG 1165

Db 903 CAAACGCGATGATTATCGACTACACCGGGATTTATAAAGCCGACATTTGGTATTAAAAATGG 962

Qy 1166 CAAATTCATGCGCATTTGGCAAGGCAAGGCAACAAAGGACATGCAAGATGGCGTAAGCCCTCA 1225

Db 963 CAAATTCATGCGCATTTGGCAAGGCGGGAACAAAGACATGCAAGATGGCGTAAGCCCTCA 1022

Qy 1226 TATGTCGTGGGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTATTACCGCTGG 1285

Db 1023 TATGTCGTGGGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTATTACCGCTGG 1082

Qy 1286 GGGAAATCGAATTCACACACCCACTTCTCTTCTCCCAACAAATTCCTTACCGCTCTAGCCAA 1345

Db 1083 GGGAAATCGAATTCACACACCCACTTCTCTTCTCCCAACAAATTCCTTACCGCTCTAGCCAA 1142

Qy 1346 TGGCGTTACAAACCATGTTTGGAGGGCGGCACAGGTCCTGTAGATGGCAAGATCGGACTAC 1405

Db 1143 TGGCGTTACAAACCATGTTTGGAGGTGGCAAGGTCCTGTAGATGGCAAGATCGGACTAC 1202

Qy 1406 TATCACTCGGGCAATGGAACCTTGCACCGCATGTTGGCGGAGCAAGAGATTTCTAT 1465

Db 1203 CATCACTCGGGCAATGGAACCTTGCACCGCATGTTGGCGGAGCAAGAGATTTCTAT 1262

Qy 1466 GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAACCTTGTAGAACAAAGT 1525

Db 1263 GAATGTAGCTTTTGGGCAAGGCAATAGTCTAGCAAAAAAACAACCTTGTAGAACAAAGT 1322

Qy 1526 AGAAGCGGCGCATTTGGTFTTTAAATTTGCATGAAGCTCGGGCAACCAAGTTCGAT 1585

Db	1323	AGAAAGCGGCGGATTTGGCTTTAAATTCATGAAGACTGGGSCACAAACACCAAGTGGCAT	1382
Qy	1586	CGATCACTGCTTGAGCGTGGCAGATGAATAAGATGTGCAAGTTGTGTATCCACACCGATAC	1645
Db	1383	CGATCACTGCTTGAGCGTGGCAGATGAATAAGATGTGCAAGTTGTGTATCCACACCGATAC	1442
Qy	1646	AGTCAATGAGCGAGGTATATGATGACACCTTAATGCAATGAACGCGGCGCCATCCA	1705
Db	1443	GGTCAATGAGCGAGGTATATGATGACACCTTAATGCAATGAACGCGGCGCCATCCA	1502
Qy	1706	TGCTTACACATTTGAGGAGCGGCTGAGGACACTCACCTGATGTATCACCATGGCAGG	1765
Db	1503	TGCTTACACATTTGAGGAGCGGCTGAGGACACTCACCTGATGTATCACCATGGCAGG	1562
Qy	1766	CGAGTCAATATTTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGT	1825
Db	1563	CGAGTCAATATTTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGT	1622
Qy	1826	TGCAGACACTTACATGCTCATGACATGACATGACACCTAGACAAACGATCCGCGAGA	1885
Db	1623	TGCAGACACTTACATGCTCATGACATGACATGACACCTAGACAAACGATCCGCGAGA	1682
Qy	1886	TTTCAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTCTCCA	1945
Db	1683	TTTCAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTCTCCA	1742
Qy	1946	TGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGAAGCAATGGGCGCTGCAGGCGA	2005
Db	1743	TGATATGGGTGTGATCGCGATGACTAGCTCCGATTCGCAAGCAATGGGCGCTGCAGG	1802
Qy	2006	AGTGATTCCTCGAATCTTGGCGAGACTGCGGATGAAGATAAAGAAATTTGGTAAGCTTCC	2065
Db	1803	AGTGATTCCTCGAATCTTGGCGAATCTGCGATGAAGATAAAGAAATTTGGTAAGCTTCC	1862
Qy	2066	TGAGATGGCAACATTAACGATATTTCCGATTAAGCGCTACATCTCCAAATACACTAT	2125
Db	1863	TGAGATGGTGCAGATTAACGATATTTCCGATTAAGCGCTATATCTCCAAATACACTAT	1922
Qy	2126	CAACCCCGCTTTGACCAACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT	2185
Db	1923	TNATCCCGCTTTGACCAACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT	1982
Qy	2186	CGCGCATTTGGTGTGTGGAATCTCGCTTTTGGCGTAAACCCCAAAATCGTGATCAA	2245
Db	1983	CGCGCATTTGGTGTGTGGAATCTCGCTTTTGGCGTAAACCCCAAAATCGTGATCAA	2042
Qy	2246	AGCGGTATGGTGTCTTCTCTGAAATGGGCGATTTCAAACGCTCTGTGCCCACTCCCA	2305
Db	2043	AGTGGCATGGTGTCTTCTCTGAAATGGGCGATTTCAAACGCTCTGTGCCCACTCCCA	2102
Qy	2306	ACCGTTTATTTACCGGAATGTTTGGGCATCACGCAAGCGGCAAAATTTGACACCGCAT	2365
Db	2103	GCCGTTTATTTACCGGAATGTTTGGGCATCACGCAAGCGGCAAAATTTGACACCGCAT	2162
Qy	2366	CACCTTTTGTTCCTCAAGTCC	2385
Db	2163	CACCTTTTGTTCCTCAAGCG	2182

RESULT 10
 AX356689
 LOCUS AX356689 2183 bp DNA linear PAT 13-FEB-2002
 DEFINITION Sequence 7 from Patent EPI1176192.
 ACCESSION AX356689
 VERSION AX356689.1 GI:18674026
 KEYWORDS
 SOURCE
 ORGANISM
 Helicobacter felis
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 Helicobacteraceae; Helicobacter.
 1
 Kusters,J.G. and Cattoli,G.
 Helicobacter felis vaccine
 TITLE

JOURNAL	Patent: EP 1176192-A 7 30-JAN-2002;
FEATURES	Akzo Nobel N.V. (NL)
source	Location/Qualifiers
	1.. 2183
	/organism="Helicobacter felis"
	/mol_type="unassigned DNA"
	/db_xref="taxon:214"
CDS	3.. 683
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23259.1"
	/db_xref="GI:18674027"
	/translation="MKLTPKEQKFLYYAGEVARRKRAKGLKLNQPEAIAYISAHIM DEARRKQTVAELMECMFLKDDVPMGVGMVFDLGVETAPDGTGLVTVNMPTEP DEHFKAGVFCGDKDIELNVGKEVTELEINVEGPKSLHVGSHFHFPEFNKALKLPRE KAYGRLDIPSGNTRLIGAGQTRKVLQIPLGGSKKVI GNMGLVNNIADERHKKHALDK AKSHGFIK"
	694..>2181
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23260.1"
	/db_xref="GI:18674028"
	/translation="MKMKQEVYNTYPTTGKVRGLDGLMAEVEHDDTYTGEELKFP GAGKTIREGMGQSNPDENTLDLIVITNAMI IDTYGIYKADIGIKNGKIHGIGKAGND MQDGVSFHMVVGTEALAGEMIITAGGIDSHTHFLSPQFPPTALANGVTVMFGGT GPDGTNATITPGKNLHRLMRAAEYSMNQVFLGKNSKKQLVEQVEAGAI GFK LHEQDTPSALDHCLSVADYDVOVCIHTDTVNEAGYVDDTLNMGNGRAIHAYHILG AGGSHSPVITMAGELNLPSTTPTIPTYINTVAEHLMDLMTCHHLDKRIREDLQFS QSRIRPGSIAAEDVLHDI GVTAMTSSDSQAMRAGEVIRPTWTADKNKKEFKLPED GADNDFRIKYISKYITINPALTHGVSEYIGSVBEKGKIALVWNPAPFVFKPKIVIK GGMVVFSEMGDSNASVPTPQPVYVREMFHGHGKAKPDTSITPRVSS"
ORIGIN	
Query Match	67.6%; Score 1948; DB 2; Length 2183;
Best Local Similarity	93.3%; Pred. No. 0;
Matches 2035; Conservative	0; Mismatches 145; Indels 0; Gaps 0;
Qy	206 GTGAAACTCACACCAAGAGCAAGAAAGTTCTTGTATATATATGCGGCGGAGTGGCT 265
Db	3 GTGAAACTCACACCAAGAGCAAGAAAGTTCTTGTATATATATGCGGCGGAGTGGCT 62
Qy	266 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT 325
Db	63 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTACATTAGT 122
Qy	326 GCCCATATTATGCAAGAGCGCGCGTGGAAAAAACCCTTGGCCAGCTTATGGAAGAG 385
Db	123 GCCCATATTATGCAAGAGCGCGCGTGGAAAAAACCCTTGGCTGAACCTTATGGAAGAA 182
Qy	386 TGCATGCACTTTTGAAGAAAGATGAAGTAAATGCGCGGGTGGTAAATATGTTCCCGAT 445
Db	183 TGTATGCACTTTTGAAGAAAGATGAGGTGATGCGCGGTGGGGAATATGTTCCCTGAT 242
Qy	446 CTAGTGTAGAAGCCACCTTTCTCATGCTAGCAAACTTGTAACTGTGAATTTGGCCCATC 505
Db	243 TTGGCGGTAGAAGCCACTTTCCCGATGGCACCACCTCGTAACCGTGAATTTGGCCCAT 302
Qy	506 GAACCATGATGAGCACTTCAAAGCGGCGGAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db	303 GAACCTGATGAACACTTTTAAAGCGCGTGAAGTGAATTTGGCTGTGTATAAGACATTGAG 362
Qy	566 CTCAATGAGGCAAGAGTAAACCGCACTTGAAGTTACTTAATGAAGGGCTTAATCTTTG 625
Db	363 CTCAACGTGGGTAAAGGAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTTAATCTTTG 422
Qy	626 CATGTGGGTAGCACTTTCCACTTCTTTGAAGCTTAAAGGCACTCAAGGCACTAAAATTCGATCGTAA 685
Db	423 CATGTGGGTAGCACTTTCCACTTCTTTGAAGCACTTGAAGTTCGATCGATCGGAA 482
Qy	686 AAAGCCTATGGCAAAACGCTAGATATTCCTCTGGCAACACGCTACGCAATGGGGCAGGA 745

Db 483 AAGCCTATGGCAACGCGCTAGATATTCCTCTCGGCAACAGCTACGCAITGGGGCAGGA 542
Qy 746 CAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAAGTGAITGGCATGAAC 805
Db 543 CAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAGTGAITGGCATGAAC 602
Qy 806 GGGCTTGTGAATACATCGCGGATGAAGCCATAAACATAAAGCGCTTGACAAGCGGAA 865
Db 603 GGGCTTGTGAATATTTGGCGACGACGCGCATAAACACAAGAGCACTAGACAAGGCAAA 662
Qy 866 TCTCAGCGATTTTCAAGTAAAGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAA 925
Db 663 TCTCAGCGATTCATCAAGTAAAGAGACTCCCATGAAATGAAAAAACAAGAGTATGTAAA 722
Qy 926 TACCTACGCGACCCACCAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCAGA 985
Db 723 CACCTACGCGACCCACCAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCAGA 782
Qy 986 AGTAGAATGACTATACCACTATGGGAAGAACTTTAAATTTGGCGCGGTAAAACTAT 1045
Db 783 AGTAGAATGACTATACCACTATGGCGAAGAGCTCAAATTTGGCGCGGTAAAACTAT 842
Qy 1046 CCGTAGGCGTATGGTCAGAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTCATCAC 1105
Db 843 CCGTAGGCGTATGGTCAGAGCAATAGCCCAAGATGAAAAACACCTTAGATTTAGTGATCAC 902
Qy 1106 TAACCGCATGATTATGACTACACCGGATTTTCAAAAGCCGACATTTGGGATTTAAAAACGG 1165
Db 903 CAACCGCATGATTATGACTACACCGGATTTTAAAGCCGACATTTGATTTAAAAATGG 962
Qy 1166 CAAATTCATGGCATTTGGCAAGCGAGGAAACAAGACATGCAAGATGCGTAAGCCCTCA 1225
Db 963 CAAATTCATGGTATTTGGCAAGCGGAGAAACAAGACATGCAAGATGCGTAAGCCCTCA 1022
Qy 1226 TATCGTGTGGTGTGGCAGAGACACTAGCAGGGAAGGTATGATTTATTCGCTGG 1285
Db 1023 TATCGTGTGGTGTGGCAGAGACACTAGCAGGGAAGGTATGATTTATTCGCTGG 1082
Qy 1286 GGGATCGATTTCACACACCCACTTCCTTTCTCCACAACAAATTCCTACCGCTAGCCAA 1345
Db 1083 GGGATCGATTTCGACACCCACTTCCTCTCTCCCAACAAATTCCTACCGCTAGCCAA 1142
Qy 1346 TGGCGTTACAAACCATTTTGGAGCGGCGACAGGTCCTGTAGATGGCAGCAAGTGGCACTAC 1405
Db 1143 TGGTGTACAAACCATTTTGGAGGTGGCAGAGTCCGGTAGATGGCAGCAATGGCAACC 1202
Qy 1406 TATCACTCCGGGCAATTTGAATCTGCACCGCATTTGGCGCAGCAGAGAGTATCTAT 1465
Db 1203 CATCACTCCGGGCAATTTGAATCTGCACCGCATTTGGCGCAGCTGAAGAGTATCTAT 1262
Qy 1466 GAATGTGGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAAGT 1525
Db 1263 GAATGTAGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAAGT 1322
Qy 1526 AGAAGCGGCGCGATTTGGTTTAAATTCATGAAGACTGGGGCACAACCAAGTGGCAT 1585
Db 1323 AGAAGCGGCGCGATTTGGTTTAAATTCATGAAGACTGGGGCACAACCAAGTGGCAT 1382
Qy 1586 CGATCACTGCTTGAAGCGTGGCAGATGAATAGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1383 CGATCACTGCTTGAAGCGTGGCAGATGAATAGATGTGCAAGTTTGTATCCACACCGATAC 1442
Qy 1646 AGTCAATGAGCGAGTTATGTAGATGACACCTTAATCAATGAACGGCGGCCATCCA 1705
Db 1443 GGTCAATGAGCGAGTTATGTAGATGACACCTTAATGAATGAACGGCGGCCATCCA 1502
Qy 1706 TGCCTACCAATTTGAGGAGGGGTGGAGGACACTCACCTGATGTATCACCATGCGCAGG 1765
Db 1503 TGCCTACCAATTTGAGGAGGGGTGGAGGACACTCACCTGATGTATCACCATGCGCAGG 1562
Qy 1766 CGAGCTCAATTTCTACCTCTCTCCACCAACCCCACTATTCCTATACCAATTAACGGT 1825
Db 1563 CGAGCTCAATTTCTACCTCTCTCCACCAACCCCACTATTCCTATACCAATTAACGGT 1622

Qy 1826 TGCAGAACACTTTAGACATGCTCATGCAATGCAACCACTTAGACAAACGCAATCGCGAGGA 1885
Db 1623 TGCAGAACACTTTAGACATGCTCATGCAATGCAACCACTTAGACAAACGCAATCGCGAGGA 1682
Qy 1886 TTTTCAATTTTCTCAAGCGGTATCCGCCCGCTCTATCGCGGTGAAGATGTGCTCCA 1945
Db 1683 TTTTCAATTTTCTCAAGCGGTATCCGCCCGGTATCTATTCGCGTGAAGATGTGCTCCA 1742
Qy 1946 TGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGGA 2005
Db 1743 TGATATGGGTGTGATCGCGATGACCTAGCTCCGATTCGCAAGCAATGGGGCGTGCAGGA 1802
Qy 2006 AGTGATTCCTCGAATCTGGCAGACTGCGGATGAAGATAAAAAAAGTATTTGGTAAAGCTTCC 2065
Db 1803 AGTGATTCCTAGAACTTGCAAACTGCAGATAAGAAATAAAAAAAGTATTTGGTAAAGCTTCC 1862
Qy 2066 TGAAGATGGCAAGATAAGATAATTCGGCATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGGTGCAGATAAGCAAACTTCGGCATCAACGCTATATCTCCAAATACACTAT 1922
Qy 2126 CAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 2185
Db 1923 TAATCCCGCTTTGACCCAGTGGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 1982
Qy 2186 CGCGACTTGTGGTGTGGAATCTCTGAAATCGGCTTTTTCGCGTAAACCCAAATCGTGATCAA 2245
Db 1983 CGCGACTTGTGGTGTGGAATCTCTGAAATCGGCTTTTTCGCGTAAACCCCTAGATTTGATTA 2042
Qy 2246 AGGCGGTATGTGGTCTTCTCGAATGCGGCAATTAACGCGTCTGTGCCACTCCCCA 2305
Db 2043 AGTGGCATGTGGTGTCTCTGAAATGGGCAATTTAAACGCGTCCGTGCCACGCGCTCA 2102
Qy 2306 ACCGGTTATTTACCCGAAATGTTTGGGCATCACGCAAGGCGAAATTTTGACACCAAGCAT 2365
Db 2103 GCGGTTTATTTACCCGAAATGTTTGGGCATCACGCAAGGCGAAATTTTGACACCAAGCAT 2162
Qy 2366 CACTTTTGTTCCTCAAGTCG 2385
Db 2163 CACTTTTGTGTCCTCAAGCG 2182

RESULT 11

AF330621
LOCUS AF330621
DEFINITION Helicobacter bizzozeronii tRNA ribosyltransferase gene, partial cds; GidB, complete cds; urease gene cluster, complete sequence and unknown gene.
ACCESSION AF330621
VERSION AF330621.1 GI:27462193
KEYWORDS
SOURCE Helicobacter bizzozeronii
ORGANISM Helicobacter bizzozeronii
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 8406)
AUTHORS Zhu, J., Teng, C.H., Chang, C.F., Chang, C.D., Simpson, K.W., Wei, C., McDonough, P., McDonough, S. and Chang, Y.F.
TITLE Cloning and characterization of a Helicobacter bizzozeronii urease gene cluster
JOURNAL DNA Seq. 13 (6), 321-331 (2002)
PUBMED 12652903
REFERENCE 2 (bases 1 to 8406)
AUTHORS Zhu, J. and Chang, Y.F.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) Population Medicine and Diagnostic Science, College of Veterinary Medicine, Cornell University, Tower Road, Ithaca, NY 14853, USA
FEATURES
source
1. 8406
/organism="Helicobacter bizzozeronii"
/mol_type="genomic DNA"
/db_xref="taxon:56877"


```

CDS
<1..929
/codon_start=3
/transl_table=11
/product="tRNA ribosyltransferase"
/protein_id="AAO15370.1"
/db_xref="GI:27462194"
/transl_table="LMVERASNKLTSTKEIFDFPPHALVVLNDTKVILKARLPAH
TAQKRPVEILLHATSTACLAQMRKVRGLVLGLEGGYSCVEVLVDNGLRVLSPK
FSGSAGHLLKNGFVITLHVGAGTFLSQSEDIRHPHAEFCRLPQAGLTDEA
PYILCVGTATLAREYKRGFGKPCBLFLHLGNPVKHAQALLTNFHLPESTLMLVA
SMVGLDKLEYLYKIAIEBHYRFYSYGDGMLIL"
926..1465
/codon_start=1
/transl_table=11
/product="GidB"
/protein_id="AAO15371.1"
/db_xref="GI:27462195"
/transl_table="MRAKLDLFTLLLEWGSVHNLGAKNHQDIEHNIQDSLOVDFI
APKSCDIDGAGFPALPQKDEADQDQYQSFVAKIPGAIAPTAASLH
FRIEHSVPSQVLVLTISRALMNAQNLIASAPLREQGHFLFYKGSHLRTEIACADH
ECHYGRVYFYSQRDFA"
1670..2053
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAO15372.1"
/db_xref="GI:27462196"
/transl_table="MLVFGHPQPCMPFKCVADIQAQVPSHTIPFFQCTHTDAFGL
SKHCSQGVYVWVQNPLESCALMANKPFTYILAENQLEACQSIANDYLPDARILKVI
EHAQLEEVUKLRDGVAFKRSFLWL"
2341..3045
/codon_start=1
/transl_table=11
/product="urease A"
/protein_id="AAO15373.1"
/db_xref="GI:27462197"
/transl_table="MKLTPKELDKMLHYAGELAKKRKANGVKLNTYEAVALLSAHVM
EEARAGKSVADLMQKRTLLKADQVMPGVAMHIEVGIANPDPGKLVITHTPVED
GKHLAPGEVLKNEITLNAQOATLEVNHGKDRPVQVSGSHVFPFVKNKLEFDR
KAYGKRLDIASCTAVRPEPEKTKVELIQGGNQIRYGFNSLVDRQADTDGKKLALR
AKHGFGVNVCGDKK"
3061..4770
/codon_start=1
/transl_table=11
/product="urease B"
/protein_id="AAO15374.1"
/db_xref="GI:27462198"
/transl_table="MKKISKEYSVSMYPTTGDKVRIGDITDILEVEHDCITYGEEIK
FGGKTRDGMQNTSPSSHELDVLNTNALIVDTYGIYKADIGIKNGKIHGIGKAGNK
DMODGVNMLCVGPATEALAEGLIVTAGGIDTHIFISPOQIPAPASGITTMIGGG
TGPDAGTNATITIPGRNWLAKMLRASEYAMNLYLGKGNVSVEPSLVDOLRAGAIGF
KIHEDNGSTPAALYHCLNLVADKYDQVAIHTDILNEAGCVEDTLQAIAGRIHTFHE
GAGGHAPDVRKWSGFNLPASTNPTIPTVTNTEAHMDMLVCHHLDRKNIKEDVQF
ADSRIRPQTAAEDKLHDMGIFSTISDSQAMRGVEVITRTWTQADKKNKGFRLPE
EKGDNDNFKIRKYSIKTYINPAIAGHSYGVSVGVEGFKADLWSPAPFGIKPNMI
GGGFIALSOMGDANASIPTPQPVYVYREMFHGHGKAKFDNIITFVSQVAYDNGIKERIG
LQRVWLPVRNCRNITKDLKFENDVTAHIEVNPETYKVKVDGKEVTSKADKISLAQLY
NLF"
4965..5552
/codon_start=1
/transl_table=11
/product="urease I"
/protein_id="AAO15375.1"
/db_xref="GI:27462199"
/transl_table="MLGLVLYYVAIVLISNGVCLANVDARSKVMNVFVGGSIICN
VIAIVSTFNPTVTCAEADVQSHLINFYGPATGCLLGFPTMYAAIINNIYNLDWK
PYGWYCLFVINTIPAILSHYSDQDLHRLGITEGDWWAFITLWANGVWLWTGWBIC
ALGKSLGKVPWLAIIEGVTITAPIWLLFIQHS"
5563..6093
/codon_start=1
/transl_table=11
/product="urease E"

```

```

/protein_id="AAO15376.1"
/db_xref="GI:27462200"
/transl_table="MLASSVLGNLKEGSSKELDFIDLEWFDQAKRMGRFTSOKGAEL
VLKKNPKMGLCDGDILFEDATSLIANNIPTPLHYVADSTAQVRLCYEYGNRHA
SLYVGDSPSLKPTPPEPRLQVLFDKLALRYEVLKSLDASQISVSAPHADPLQEGSA
PLKFKSALUDLOIVIKK"
6098..6775
/codon_start=1
/transl_table=11
/product="urease F"
/protein_id="AAO15377.1"
/db_xref="GI:27462201"
/transl_table="MQKDFLLQVNDAMFPIGSYTHSFCLETYIOHKEYSSKESALDY
MRAVSTOFLYTELLALAYTAHQEDLESLENEEQICLATPPELRSANQKLGNR
FLKTIGVLDPLSAFFKDYMQKSTPTTHASAYFCACGLGLEASLKHLYLAQSNM
VINCVKTIPLAQNDQRIALLAQETFTTLLVTLEGLDASYLCAASIQNDIKAMQHEHL
YSRLYMS"
6797..7396
/codon_start=1
/transl_table=11
/product="urease G"
/protein_id="AAO15378.1"
/db_xref="GI:27462202"
/transl_table="MWKIGVCGPVSGKGTALIEALTRAMSQOYSIGVITNDIYTKEDA
EFLCRNSVMPRERIGVETGCPHTAIRDASMNLEAVEELHAKFPDIELMFIESGD
NLSTTFNPELADFTIFVIDVAEGDKIPRKGGGPIRSDLLIINKIDLAPYVGADLGM
DRDSKMRGDKPFLFTNIRSKGLDHVIGWKYKLYALLED"
7399..8232
/codon_start=1
/transl_table=11
/product="urease H"
/protein_id="AAO15379.1"
/db_xref="GI:27462203"
/transl_table="MAVOHPTSYAQAPRLYLRKTIQONGKVCVIADNYFSPPEKLMPPF
YERGNMAEILIIAVSPGMLKGAQOIQMDIGENCOLELSFQSFKEIQDTEDEGVATNT
QIHHPNALLDFSPPLPIPPANAHFKNHSIHLVHKNQALLYSEIITAGHIANGEAFAP
HKHSTUKISTQEGESRVSFNDITILEPASMDLKNPCWFGNYTHYIFLRLDKL
EDLLELLESTQINAGVSLTLRLPEGYHSLCLKALANGSEPLLALURKSISNLLVERME
G"

```

ORIGIN

Query Match	Best Local Similarity	Score	DB	Length
39.4%	67.2%	1135.6	DB 15	8406
Matches 1723	Conservative 0	Mismatches 809	Indels 32	Gaps 2
Qy	73	TATTATTAAATTTTAAATAATTTACTTATCATATATAATAATATTTACTTTATTTATTA 132		
Db	2208	TAAGATCAACTAGATAATAATAATATTTATCTTGTCTAGAAATAATATTAACAAGGTAGC 2267		
Qy	133	AAAAAGTTAAATAAAAGTAACGAAATTAGGACTATAATCCCATTTGCTTTTAAATAATTAA 192		
Db	2268	CAAAACAGACTAGAAATTTGTCCTGATAGCTATTCAATAATACAAATTTGGTA 2327		
Qy	193	CAAGAGTAATAGGTGAACTCACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252		
Db	2328	GAAGGAGTTTAGGATGAAATTAACCCCTTAAAGAGCTGCAAGCTCATGTTGTCATTATGC 2387		
Qy	253	GGGCGAAGTGGCTAGAACGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAGGACCAT 312		
Db	2388	GGCGCAATTTGGCTAAAAAACGCAAAATGGCGTTAAAGCTTAAATATATCTAGGACAT 2447		
Qy	313	TGCTTTACATTAGTGCCTCATATTATGACGAGAGCGCGCTGGAAAAAACCCTTTGCCCA 372		
Db	2448	AGCCTCATAGTGCCTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2507		
Qy	373	GCTTATGGAAGAGTGCACTGCACTTTTGTGAAAAAGATGAAGTAATGCCCGGGTGGTAA 432		
Db	2508	TTTGATCAAGAGGAGGACACTTCTTAAAGCTGATGATGATGATGATGATGATGATGATG 2567		
Qy	433	TATGTTCCCGATCTAGGTGTAGAGCCACTTCTTCTGATGTTAGTGTAGTGTAGTGTAGT 492		
Db	2568	TATGATCAACGAAAGTGGGATTTGAAGCTTAACCTTCCCTGATGGGACAAACCTGTAACCAT 2627		

QY 493 GAATTTGGCCCATCGAACCAGATGAGCACTTCAAAGCGGGCGAAGTGAATTTGGTTGCGA 552
Db 2628 CCATACCCCGCTGAAGATGTGGGCAATAAATTTGGCTCCGGGTGAAGTGAATTTGAAAAA 2687
QY 553 TAAAGACATCGAGCTCAATGCAAGCAAGAAAGTAAACCGAATCTGAGGTTTACTAATGAAGG 612
Db 2688 CGAAGACATCACTTTGAATGCAAGCAAGAAAGTAAACCGAATCTGAGGTTTACTAATGAAGG 2747
QY 613 GCCTAAATCTTGTGATGAGGTTAGCAATTTCCACTTTCTTTGAAGCTAAACAGGCACTAAA 672
Db 2748 CGATCGCCCGCTGCAAGTGGGCTCCCACTTCCACTTTCTTTGAAGTGAATGAAGCTTTTGA 2807
QY 673 ATTGATCGTGAAGAGCTATGCAAGAGCTAGATATTCCCTCTGCAACACCGCTAGG 732
Db 2808 ATTTGATCGTGAAGAGCTATGCAAGAGCTAGATATTCCCTCTGCAACACCGCTAGG 2867
QY 733 CATTTGGGCGAGCAAAACCCGAAAGTGCAGTTGATTCCTCTGCTGGCGAGTAAAAAGT 792
Db 2868 CTTTGAACCCGCTGGAAGAAACCGTGGAAATTTGAATTTGGCGGTAAACCAAGCAT 2927
QY 793 GATTGGCATGAACGGGCTTTGTAATAACATCGCGGATGAACGCCATAAACAATAAAGCGCT 852
Db 2928 TTACGGCTTTAACTCTCTTTGTTGGATCGCAAGCGGCTAGACATTTCTTCTGCAAAAACCTTGTCT 2987
QY 853 TGACAGCGGAAATCTCAGGATTT-----ATCAAG 883
Db 2988 CAAACGCGCCAAAGAACATGGCTTTTGGTGTGTTGAAATTTGCGGTTGCGATFAAAAAATAAGG 3047
QY 884 TAAAGAGACTCCCATGAAA---ATGABAAAACAAGAAATGTAATATCTACGGACCCAC 940
Db 3048 AAAGGACAAATCCGATGAAAAAATACTCGAAAAAGAAATGTTTCTATGATGGAGCCAC 3107
QY 941 CAAAGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTTGGGAGAGTAGAATGAAACATGACTA 1000
Db 3108 TACGGCGATAAAGTGAATTTGGCGATACCGACTGATCTTAGAAGTCGAACATGACTG 3167
QY 1001 TACCCTATGCGGAAGAACTTAAATTTGGCGGGGTAAACATATCCGTGAGGGTATGG 1060
Db 3168 CACCACCTATGCGGAAGAAATTAAGTTTGGTGGCGTAAAAACCATTCGCGATGGGATGCG 3227
QY 1061 TCAGAGCAATAGCCTGATGAACACACCTAGATTTAGTCACTACCTAACGCGATGATTAT 1120
Db 3228 ACAACCAACAGCCCCAGCAGCACTCGATCTTGTGCTCACTAACGCCCTGATCGT 3287
QY 1121 CGACTACACCGGATTTTAAAGCCGACATTTGGGATTAAGAAACGGCAAAATCCATGGCAT 1180
Db 3288 GGAATTACACCGCATTTTAAAGCCGATTTGGCAATTAAGAAATGCAAAATCCATGGCAT 3347
QY 1181 TGGCAAGGCAGAAACAAGGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGTGT 1240
Db 3348 TGGCAAGGCAGCAATAAAGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGTGT 3407
QY 1241 GGGCACAAGACATGACAGGGAAAGTATGATTTATACCGCTGGGGGATCGATTCA 1300
Db 3408 TGCTACTAGGGCTTTGGCGCTGAAAGGGCTGATTTGTTACAGCTGGTGGGATTTGACACCCA 3467
QY 1301 CACCACCTTCTTCTCCACAACAATTTCCCTACCGCTCTAGCCCAATGGGTTTACAACCAT 1360
Db 3468 CATCCACTTTATTTCTCCCAACAATTTCCCAAGATTTGGCAGGATCACACCAT 3527
QY 1361 GTTTGGAGCGGCACAGGCTCTGTAGATGGCAAGATGCGACTACTATCACTCCCGGCA 1420
Db 3528 GATTGTTGGGGAACAGGTCAGCTGATGGGACTAACGCGACTACCATCACTCCCGGGCG 3587
QY 1421 ATGGAATTTGCAACCGCATTTGGCGGCGAGGAGATTTCTATGAATGTGGGCTTTT 1480
Db 3588 CTGGAACCTTTAAACCAATGCTCCGTGCTCTGAAGATATGCAATTTGCGGCTATTT 3647
QY 1481 GGGCAAGGCAATAGCTTAGCAAAAACAACTTTGTAAGCAAGTAGAAGCGGGCGCAT 1540
Db 3648 GGGTAAAGGGAATGTGCTTTATGAACCCCTCCCTGGTGGATCAACTCGAAGCTGGAGCCAT 3707
QY 1541 TGGTTTTAAATTGCAAGACTGGGGCACACACCAAGTGGCATGATCACTGCTTGTAG 1600

RESULT 12

Db 3708 TGGCTTTAAAAATCCAGAACTGGGGTAGCACACCTGCAGGCATCTACATTTGCTTGAA 3767
QY 1601 CQTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGGCAAG 1660
Db 3768 TGTGGCTGCAAAATACGATGTGCAAGTGGCTATCCACACCGGATACCTTGAATGAGGGG 3827
QY 1661 TTATGTAGATGACACCTTAAATGCAATGAACGGGGCGGCTATCCATGCGCTACACATTTGA 1720
Db 3828 CTGTGTGGAAGACACTTTTGCAAGCCATTTGCTGGGGGCACTATCCACACTTTCCACACTGA 3887
QY 1721 GGGAGCGGTGAGGACACTCACCCTGATGTTATCACCATGGCAGGCGAGCTCAATTTCT 1780
Db 3888 AGGTGCTGGTGGGGGCGCTCCCGATGTCAATGAATGTCTGGCGAATTTTAAACATCCT 3947
QY 1781 ACCCTCTCCACACACCCCACTATTCCCTATATACCAATTAATAGGTTGCGAAGACACTTAGA 1840
Db 3948 CCCAGCTTCTACCAACCCCACTTCTTTCACCGGTGATACAGAAGCCGAACACATGGA 4007
QY 1841 CATGCTCATGACATGCAACCTAGACAAACGCAATCCGCGAGGATTTTACAAATTTTCTCA 1900
Db 4008 CATGTTGATGTGTGCCACACTTGGATAAAAACATCAAGAAGATGTCCAGTTTGTCTGA 4067
QY 1901 AAGCCGTATCCGCCCGGCTCTATCGCGCTGAAGTGTCTCCATGATATGGTGTGAT 1960
Db 4068 TTCTAGGATTCGCCCCCAACCATCGCGCTGAGGACAACTCCAGATATGGGATTTT 4127
QY 1961 CCGCATGACAAGCTCGGATTCGAAGCAATGGGGGTGCGAGCGCAAGTATTCCTCGAAC 2020
Db 4128 CTCTATCACAGCTCTGACTCCCAAGCGATGGCGGTGTAGCGGAGTCAATCACCCGAC 4187
QY 2021 TTGGCAGACTGCGGATAAGAATAAAAAAGAAATTTGGTAACTTCTGTAAGATGGCAAGA 2080
Db 4188 TTGGCAAAACAGCGGACAAAAAAGAAATTTGGTCTGCTGCGCTGAGGAAAAAGCGGA 4247
QY 2081 TAAACGATTAATTTCCGATTTAAGCGCTAGATCTCCAAATACATCTCAACCCCGCTTGAC 2140
Db 4248 TAATGACAACTTTCCGATCAAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTGC 4307
QY 2141 CCACGCGGTGACGAGTATATCGGCTCTGTGAAGAGGCAAGATCGCGCACTTTGGTGTG 2200
Db 4308 ACACGCGATTTCTGAATATGTCGCTCTGTAGAAGTGGGCAAAATTCGCCGATTTGGTCT 4367
QY 2201 GTGGAACTCTGCTTTTGGGCTAAAAACCCAAATCGTGATCAAAAGCGGATATGGTGT 2260
Db 4368 TTGGAGTCTCTGCTTTTGGCATTAACCCCAACATGATCATCAAAAGCGGATTCATCGC 4427
QY 2261 CTTCTCTGAATGGGCGATTTAAACGCTCTGTGCCACTCCCAACCGGTTTATTACCG 2320
Db 4428 ACTTTCTCAAATGGGCGATGCCAATGCTCTATCCCACTCCCAACCCCGTGTATTACCG 4487
QY 2321 CGAAATGTTTGGGCACTACGCGCAAGCGAAATTTTGACACCGATCACTTTTGTTCCAA 2380
Db 4488 CGAAATGTTTGGCCCATGTTGAAGCCAAATTTGACCAATATCACATTTGTATCCCA 4547
QY 2381 AGTCGCTATGAAAATGGCGTGAAGAAAGCTGGGCTTTAGAGCGCAAGTCTTACCGGT 2440
Db 4548 AGTGGCTTATGACACCGCATTTAAGAAGATTTGGCTTTGCAAGAGTGGTTTGGCAGT 4607
QY 2441 CAAAACTCCCGTAAACATCAACGAAGAACTTTCAAGTTCAACGACAAACGCGCAAAAT 2500
Db 4608 TAAAACTCCCGCAACATCAACAAAAAGACCTCAAAATCAACGATGTTTACCGCACAT 4667
QY 2501 CACCGTCAATCCGAACCTTCGAGTCTTTGTAGATGCAAACTCTGCAACCTCTTAAAC 2560
Db 4668 CGAAGTCAATCTGAAACCTTACAAAGTTAAAGTGGATGGCAAGAGTTTACTTCCAAAGC 4727
QY 2561 CACCTCGCAAGTGCCTCTAGCCCGAGCGCTACACTTTTCTTCTAGG 2604
Db 4728 AGCGGATAAATCAGCCTAGCACAACTCTACAACTTTGTTCTAGG 4771

HECUREASE 2664 bp DNA linear BCT 03-MAY-1995
 Locus HEcurease helicmannii urease, complete cds's.
 DEFINITION
 ACCESSION L25079
 VERSION L25079.1 GI:793908
 KEYWORDS urease.
 SOURCE Candidatus Helicobacter heilmannii
 ORGANISM Candidatus Helicobacter heilmannii
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.
 REFERENCE 1 (bases 1 to 2664)
 AUTHORS Solnick, J.V., O'Rourke, J., Lee, A. and Tompkins, L.S.
 TITLE Molecular analysis of urease genes from a newly identified uncultured species of Helicobacter
 JOURNAL Infect. Immun. 62 (5), 1631-1638 (1994)
 PUBMED 8168924
 COMMENT On May 3, 1995 this sequence version replaced gi:529422.
 Original source text: Helicobacter heilmannii (individual isolate 2) DNA.

FEATURES
 source
 Location/Qualifiers
 1..2664
 /organism="Candidatus Helicobacter heilmannii"
 /mol_type="genomic DNA"
 /isolate="2"
 /db_xref="taxon:35817"
 /lab_host="BALB/c"
 200..204
 211..915
 /codon_start=1
 /transl_table=11
 /product="urease"
 /protein_id="AAA65722.1"
 /db_xref="GI:529423"
 /translation="MKLTPELDDKMLHYAGELAKORKAKGIKLNYTEAVALISAHV
 EEARAKKSVADLMQSGRILLKADDDVMPGVAMHIEHVGIEAGPDDGTLVITHPVEA
 GSDKAPGVILKNNEDITNAGKHAVOLKKNKGRDPVQVGSFHFEFVYNNKLLDFDRE
 KAYGKRLDASGTAVRPFPECEETVELIDIGGNKRIYGFNALVDRQADHDGKKLAKR
 AKEKHFGTINCGCDNR"
 919..923
 930..2636
 /codon_start=1
 /transl_table=11
 /product="urease"
 /protein_id="AAA65723.1"
 /db_xref="GI:529424"
 /translation="MKKISKEVSVMYGPTTGDKVRLGDTDLILEVEHDDCTTYGRIK
 FGGKTRIDMGMTNPSPSHEDLDVITNALIVDYGTGYKADIGIKNKGKIHGIGKAGN
 DLQDGVNRLCVGPTAAKLAAGLIVTAGGIDITHFIISPOQIPTAFASGITTTGGG
 TPDAGTNAITTPGRWNLEAMESEBYAMNLGYLKGNVSPFALIDQLBAGAI
 KIHEDGSTPSALNHALNADKYDVQVAIHTDITLNEAGCVEDTLEAIGRTITHTHE
 GAGGCHADVPINWAGFNILPASTNPTIPTKNTAEHMDMLMCHLDKNIKEDVEFA
 DSRIRPQTAAEKLHDMGIFSITSSDSQAMGRVSEVITRTQTDKVKKEGFRLEPE
 KGNDNFRILKRYISKYITINPAITHGLISEYVGVGVADLVLSWPAFGIKENMIIL
 GGFIALSQMGDANASITPPQVYTYREMFGHGKAKFDPTNITFVSQVAVENGIKHELJ
 QRVLVLPKNCNRNITTKDLKFENDVTAHIEVNPETYKVKVDGNEVTSHAADKLSLAQLN
 LF"

ORIGIN

Query Match 38.2%; Score 1099.8; DB 15; Length 2664;
 Best Local Similarity 67.0%; Pred. No. 0;
 Matches 1722; Conservative 1; Mismatches 813; Indels 36; Gaps 9;

QY 63 TTGTTAATCTATTATTTTTHAATAATTAATCTATTATCATATATAATAATATTATT 122
 DB 70 TTGTGAATAGATTAACAAGAATAAATAATATATCTTTTGATA-ACAAAATATTAACA 128
 QY 123 ACTTATATTTAAAAGTTAATAAAAAGTAACGAAATTAGGACTATAATCCCATTCGCTTTA 182
 DB 129 ACCATTAGCCAAAGATACATAAAATCTGCCTGTTGATAGCT-TGGCGTATTCAATAACA 187
 QY 183 AAATTTAACACAGGAGTAAATAGGTGAAACTCACACCCAAAGACGAAAGAAAGTTCTTTGT 242
 DB 188 CCATTTTATACGAGGAGTTTAGATGAACACTGACACCTAAAGAGTTGGATAAGTTGATGTC 247

1292 CGATTTCACACACCCTTCCTTCTCCACAAATTCCTTACCGCTCTAGCCATGGCGT 1351
1328 CGACACCCACATCCATCTTCTCTCCAAATCCCTACTGCTTGTAGCGGAAT 1387
1352 TACAAACATGTTGGAGCGGCACAGGTCCTGTAGATGGCAGCAATGGACTTACTATCAC 1411
1388 CACACCATGATCGGGGGGGGACTGACCTGACATGGACGACCAACGACCATCATCAC 1447
1412 TCCGGGCAATGGAATTTGCACCGCATGTTGCGGACGAGCAGAGAGTATTTCTGAATGT 1471
1448 TCCGGCGCGCTGGAATTTAAAGAAATGCTCGCGCTTCTGAAGAAATAGCCATGAACCT 1507
1472 GGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAACAACTTTAGTAAACAAAGTAGAAGC 1531
1508 TGGCTACCTTGGTAAAGGGGAATGTCTTTTGAACCTGCTCTCATTTGACCAAGCTAGAACG 1567
1532 GGGCGCATTTGTTTAAATTCATGAAGACTGGGGCAACACCAAGTCGATCATCA 1591
1568 CGGCGCATTTGGCTTTTAAATTCAGAAAGACTGGGGCAGCAACCTCAGGCATCAACCA 1627
1592 CTGCTTGAGCGTGCAGATGAATCGATGTGCAAGTTTGTATCCACACCGGATCAGTCAA 1651
1628 CGCTCTAAACATCGCTGACAAATACGATGTGCAAGTGGCGATCCACACCGACACCTTGAA 1687
1652 TGAGGCGAGTTATGTAGATGACACCTTAAATGCAATGAACCGGGCGGATCCATGCTCTA 1711
1688 TGAAGCGGCTGTGTGAAGACACCTTGAAGCCATCGCTGGAGCAGCAATCCACACCTT 1747
1712 CCACATTTGAGGCGGGTGGAGGACACTCACTGATGTTATCCATCCATGGCAGGAGCT 1771
1748 CCACACGGAAGTGTGGCGGGGACACGCTCTGACGTGATCAAAATGGGCGGGAAT 1807
1772 CAATATTTACCTCTCCACACCCCACTATTTCCCTATACCAATTAATACGTTTGAGA 1831
1808 TAAATCTCTCTGTTCTACCAACCCCACTATTTCCCTTCCACCAAAACACAGAGCCGA 1867
1832 ACATTTAGACATGTCATGACATGCCACCACTAGACAAAACGATCCGCGAGGATTTACA 1891
1868 ACATGACATGATCTTATG---TGCCACCACTTGGATAAAAACATCAAAAGAGTGTGA 1924
1892 ATTTTCTCAAGCCGATATCCGCGCGGCTCTATCGCGCTGAAGTGTCTCCATGATAT 1951
1925 ATTTGCGGACTTCACGATTCGCGCGGCAACCAATTCGCGGGAAGATAAATCCACGACAT 1984
1952 GGGTGTGATCCGATGACAGCTCGATTCGCAAGCAATGGGCGTGCAGCGCAAGTGTAT 2011
1985 GGGGATTTCTCCATCACTAGCTCGACTCTCAAGCCATGGTCCGCTGGGCGGAAGTGTAT 2044
2012 TCCTCGAACTTTGGCAGACTGCGGATTAAGAAATTAAGAAATTTGGTAAAGCTTCCTGAAGA 2071
2045 CACCGGCACTTTGGCAACAGCGGACAAAACAAAAAGAAATTTGGCGCTTGGCTGAAGA 2104
2072 TGGCAAGATTAAGATTAATTTCCGATTAAGCGGTATCTCTCAAAATACATATCAACCC 2131
2105 AAAAGGCGACACACACATTTCCGATCAAAACGCTATATTTCCAAATACACCATCAACCC 2164
2132 CGCTTTGACCCACGCGTGGAGGATATATCGGCTCTGTGGAAGGCGCAAGATCGCGCA 2191
2165 TGCCATCACACGCGCATTTCTGAATATGTCTGCGCTCTGTAGAAGTGGGTAATACGCTGA 2224
2192 CTTGCTGTGTGAATCTCTGCTTTTGGGCTTAAACCCAAATTCGTGATCAAGGCGG 2251
2225 CTTGCTGTGTGGAGCCCTGCTTCTTTGGCATTAACCCCAACATGATCATCAAGGCGG 2284
2252 TATGCTGTCTTCTGAAATGGGCGATTTCAACGCGCTGTGTGCGCATCTCCCAACCGGT 2311
2285 TTTTATTTGCGCTTCTCAAAATGGGCGATGCGCAACGCTTCTATCTCCCACTCTCCCAACCGGT 2344
2312 TTATTTACCGGAATTTTGGCATCGGCAAGCGGAATTTGACCCAGCATCATCTTT 2371
2345 GTACTACCGGGAATTTGCGGCCACCAACGCGCAAGCCAAATTTTGACCAACATCATCTTT 2404
2372 TGTTTCCAAAGTTCGCTATGAAATGGGCTGAAAGAAAAGCTGGGCTTTAGAGCGGCAAGT 2431

Db 2405 CGTGTCTCAAGTGGCTTATGAAAAACGGCATTAACACAGATGGGCTTCAAAGAGTTGT 2464
Qy 2432 TCTACCGGTCAAAAACTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACACGACAAAAC 2491
Db 2465 GTTGCTGTGAAAAACTGCCGCAACATCACCAAAAAAGACCTTAAGTTCAACGATGTCAC 2524
Qy 2492 GCGAAAAATCAACCGTCGATCCGAAACCTTCGAGGCTTTTGTAGATGGCAAACTCTGCAC 2551
Db 2525 CGCACATATCGAAGTCAACCTTGAACCTTACAAAGTGAAGTGGATGGCAACGGAAGTTAC 2584
Qy 2552 CTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTTCTTCTAG 2603
Db 2585 CTCCTCATCGGCTGACAAATTTAGCGCTAGCACAACTCTATAACCTATTCTAG 2636

RESULT 13
AE001439_00/c
WPCOMMENT

Sequence split into 17 fragments LOCUS AE001439 Accession AE001439

Fragment Name	Begin	End
AE001439_00	1	110000
AE001439_01	100001	210000
AE001439_02	200001	310000
AE001439_03	300001	410000
AE001439_04	400001	510000
AE001439_05	500001	610000
AE001439_06	600001	710000
AE001439_07	700001	810000
AE001439_08	800001	910000
AE001439_09	900001	1010000
AE001439_10	1000001	1110000
AE001439_11	1100001	1210000
AE001439_12	1200001	1310000
AE001439_13	1300001	1410000
AE001439_14	1400001	1510000
AE001439_15	1500001	1610000
AE001439_16	1600001	1643831

LOCUS AE001439 1643831 bp DNA circular BCT 19-JAN-2006

DEFINITION Helicobacter pylori J99, complete genome.

ACCESSION AE001439 AE001440-AE001571

VERSION AE001439.1 GI:12057207

KEYWORDS

SOURCE Helicobacter pylori J99

ORGANISM Helicobacter pylori J99

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.

REFERENCE 1 (bases 1 to 1643831)
Alm, R.A., Ling, L.-S.L., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G., Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C., Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E., Vovis, G.F. and Trust, T.J.

TITLE Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori

JOURNAL Nature 397 (6715), 176-180 (1999)

REFERENCE 2 (bases 1 to 1643831)
King, B.L., Alm, R.A. and Trust, T.J.

AUTHORS Direct Submission

TITLE Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney Street, Cambridge, MA 02139, USA

JOURNAL On or before Jan 19, 2006 this sequence version replaced

COMMENT gi:4154501, gi:4154511, gi:4154523, gi:4154537, gi:4154549, gi:4154559, gi:4154573, gi:4154583, gi:4154594, gi:4154604, gi:4154617, gi:4154626, gi:4154639, gi:4154651, gi:4154666, gi:4154678, gi:4154689, gi:4154703, gi:4154713, gi:4154723, gi:4154733, gi:4154749, gi:4154760, gi:4154775, gi:4154789, gi:4154799, gi:4154813, gi:4154826, gi:4154838, gi:4154861, gi:4154869, gi:4154880, gi:4154897, gi:4154910, gi:4154929, gi:4154939, gi:4154949, gi:4154961, gi:4154971, gi:4154979, gi:4154992, gi:4155005, gi:4155018, gi:4155034, gi:4155043, gi:4155057, gi:4155074, gi:4155086, gi:4155100, gi:4155102,

gi:4155115, gi:4155127, gi:4155147, gi:4155161, gi:4155170, gi:4155179, gi:4155191, gi:4155199, gi:4155211, gi:4155220, gi:4155238, gi:4155252, gi:4155264, gi:4155275, gi:4155285, gi:4155295, gi:4155310, gi:4155326, gi:4155338, gi:4155350, gi:4155371, gi:4155382, gi:4155392, gi:4155407, gi:4155418, gi:4155425, gi:4155431, gi:4155444, gi:4155454, gi:4155471, gi:4155483, gi:4155492, gi:4155505, gi:4155513, gi:4155523, gi:4155533, gi:4155550, gi:4155558, gi:4155575, gi:4155590, gi:4155602, gi:4155617, gi:4155628, gi:4155636, gi:4155656, gi:4155664, gi:4155675, gi:4155687, gi:4155697, gi:4155706, gi:4155717, gi:4155724, gi:4155739, gi:4155753, gi:4155763, gi:4155776, gi:4155799, gi:4155811, gi:4155845, gi:4155858, gi:4155872, gi:4155887, gi:4155898, gi:4155906, gi:4155914, gi:4155929, gi:4155938, gi:4155952, gi:4155964, gi:4155974, gi:4156044, gi:4156000, gi:4156017, gi:4156027, gi:4156032, gi:4156044, gi:4156051, gi:4156065, gi:4156083, gi:4156095, gi:4156108, gi:4156120.
 Addresses all correspondence to: hp@rcb.us.astro.com or Richard A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith, Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills, Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and Diane E. Taylor are with the University of Alberta Department of Medical Microbiology and Immunology, Edmonton, Alberta, Canada, T6G 2H7 and the Canadian Bacterial Diseases Network. All other authors are with Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Putative identifications, sequence alignments, and name and sequence search capability are available at ARCB's World Wide Web site. (URL: <http://www.astro-boston.com/hpylori/>.)
 Location/Qualifiers
 1. 1643831
 /organism="Helicobacter pylori J99"
 /mol_type="genomic DNA"
 /strain="J99"
 /db_xref="taxon:85963"
 complement(51..467)
 /gene="nusB"
 /locus_tag="jhp_0001"
 complement(51..467)
 /gene="nusB"
 /locus_tag="jhp_0001"
 /note="similar to H. pylori 26695 gene HP0001"
 /codon_start=1
 /transl_table=11
 /product="TRANSCRIPTION TERMINATION"
 /protein_id="AAD05585.1"
 /db_xref="GI:4154504"
 /translations="MATRTQARGAVVELLYAFESGNEEIKKIASMLLEKKIKNNQLA
 FALSIFNGVLEKINEIDALIEPHLKQWDFKRLGSMKAILRLGAYEIGFTPTQNP111
 NECIELGKLYAEPNTPKFLNALDLSLKKLAQKPLI"
 complement(469..939)
 /gene="ribE"
 /locus_tag="jhp_0002"
 complement(469..939)
 /gene="ribE"
 /locus_tag="jhp_0002"
 /note="similar to H. pylori 26695 gene HP0002"
 /codon_start=1
 /transl_table=11
 /product="RIBOFLAVIN SYNTHASE BETA CHAIN"
 /protein_id="AAD05586.1"
 /db_xref="GI:4154505"
 /translations="MQIEGKQLQNEKIALITSRNHIITRLKEGAMDCFKRHGG
 DEELLDLVPGAYELDLIDLSKGYDGVCLGAIIRGGPHFDYVSAEATKGIA
 SAMLKYSMPVSFGVLTNDIEQAIERAGSKAGNKGFEAMSTLIELSLCQTLDK"
 complement(949..1779)
 /gene="kdsA"
 /locus_tag="jhp_0003"
 complement(949..1779)
 /gene="kdsA"
 /locus_tag="jhp_0003"

FEATURES

source

/note="similar to H. pylori 26695 gene HP0003"
 /codon_start=1
 /transl_table=11
 /product="3'-deoxy-D-manno-octulosonic acid 8-phosphate
 synthase"
 /protein_id="AAD05587.1"
 /db_xref="GI:4154506"
 /translations="MKTENTKTPKPVLIAGPCVIESLENLRSTAIKLOPLANNERLDLF
 YPKASFQKANTRTLSYSGPGLKGLKLEMTQTIKDEFGYKILTDVHESYQASVAAKVAD
 ILQIPAFRCQRTDLIVSYQTNAINIKKGQPMNPKMQYSVLKALTKTRDSSIQSPY
 ETALKNGWLCERGSSFGYGNLVDMRSIKIMREFAPVIFDATHSVQMPGANGKSSG
 DSSPPPILPRAAAGVIGIDGLPAETHIDPKNALSDGANMLKPDLEHLVTDMLKIQNLF
 ."
 complement(1766..2431)
 /gene="icfA"
 /locus_tag="jhp_0004"
 complement(1766..2431)
 /gene="icfA"
 /locus_tag="jhp_0004"
 /note="similar to H. pylori 26695 gene HP0004"
 /codon_start=1
 /transl_table=11
 /product="Carbonic anhydrase"
 /protein_id="AAD05588.1"
 /db_xref="GI:4154507"
 /translations="MEAFIAGLEPQNEEYEPKELYESLTKQKPHTLFISCVDSRVV
 PNLTGTQPGELYIRNNGVNPPTKSYKESLSTIASVEYIAHGVQNLIIICGSDC
 GACSIHLIHDETTKAKTPYANMTQFLPEIKELKNHPQSNHFAKRSWLTERRLAR
 LQANLLSYDFIQERVINNELKIPGMHYIETGRINYNFESHFPETIETIKQRIH
 ENF"
 2552..3235
 /gene="pyrF"
 /locus_tag="jhp_0005"
 2552..3235
 /gene="pyrF"
 /locus_tag="jhp_0005"
 /note="similar to H. pylori 26695 gene HP0005"
 /codon_start=1
 /transl_table=11
 /product="OROTIDINE 5'-PHOSPHATE DECARBOXYLASE"
 /protein_id="AAD05589.1"
 /db_xref="GI:4154508"
 /translations="MQLCVALDLEKEDNLSLLQELKGLDLWAKVGLRSFIRDGATFL
 DEIRKIDGNFKIFLDKLYDIPYMANAALECAKLDIDMLTVHLSSAKALTILMORL
 NALKRPLIMGVSALTSPSEBEFLMVMYNAPLTKQAITLSAIGKESIDGVGVSFESL
 AIKEALGKGLTLPGLRDLKNDKQSERVANAKQNLSDFIYVVGRIYQAKEPRE
 VLELLKDC"
 3235..4065
 /gene="panC"
 /locus_tag="jhp_0006"
 3235..4065
 /gene="panC"
 /locus_tag="jhp_0006"
 /note="similar to H. pylori 26695 gene HP0006"
 /codon_start=1
 /transl_table=11
 /product="PANTOATE--BETA-ALANINE LIGASE"
 /protein_id="AAD05590.1"
 /db_xref="GI:4154509"
 /translations="MRALETIATLRBYRKSLSKESVGVPTMGALHRHQSILTERSKE
 NSHTIVSVFNPTQGANEDFSAYRPLEKDILALCEKSGVNAVFAPKIGMYPYEAQ
 RLKLYAPFLSHLEGAVRKGFVGVVLRFLHNTPTTRAYFGKDAQLLIIQHL
 VQDLILDIEIAPCEIVRSDNLSSRNVCINATKOKALIPKALEKIQQAIDREK
 ACEKLLGLLEILKNLEVDYLECCNHKLEPLKTIEPANTLVLAARVGVKTRLLDNLWV
 ."
 complement(5436..7443)
 /locus_tag="jhp_0007"
 complement(join(5436..7413,7415..7443))
 /locus_tag="jhp_0007"
 /note="similar to TIGR HP0009; expression of this protein
 is predicted to be regulated by slipped-strand repair at
 the (CT) dinucleotide repeat; length of the repeat has

Query Match		36.1%; Score 1040.8; DB 15; Length 110000;
Best Local Similarity		65.3%; Pred. No. 2.9e-313;
Matches 1681; Conservative		1; Mismatches 853; Indels 39; Gaps 4;
Qy	67	TAATRCCTATTATTAATTTTAAATAATTTACTT-ATTATCATATATAATAATATTACT 125
Db	75600	TAATACCTAATTTAAATAAATAATTTAGTTTAAATAACGGCTTCGTAAATCTTAGTAAA 75541
Qy	126	TATATTTAAAAAGTTAAATAAAAGTAACGAAATTTAGGACTATAATCCCAATTGCCTTTAAAA 185
Db	75540	TCAAAGATTGCTACAATTTACTCCAACTTGAATGGCTTATGCTTCAAGGAAAAACAC 75481
Qy	186	TTTAAACACAGAGTAATAGGTGAACTCACACCCAAAGAGCAAGAAAGTTCTTGTTAT 245
Db	75480	TTTAAAGATAGGAAATGAGATGAACTCACCCCAAAAGAGTTAGATAGTTGATGCTCC 75421
Qy	246	ATTATGGCGGAGTGGCTAGAAAGCCGAAGCAGAGGGCTTAAAGCTCAACCAACCG 305
Db	75420	ACTACGCTGGAGAAATTAGCTAGGAACGCAAGAAAAAGGCATTAAAGCTTAACTATGG 75361
Qy	306	AAGCCATTGCTTACATTAGTCCCATATTATGGACGAAGCGCGCTGGAAAAAAGAACCG 365
Db	75360	AAGCGTAGCTTGTATTAGTCCCATATTATGGAAAGAGCGAGCTGGTAAAAAGACTG 75301
Qy	366	TGCCCCAGCTTATGAAGAGTGCAATGCACTTTTGAAGAAAAAGATGAAGTAATGCCCGGG 425
Db	75300	CGGCTGAATTTGATGCAAGAGGGCGCACTCTTTTAAAAACCGGATGATGTGATGGTGTG 75241
Qy	426	TGGGTAAATATGGTTCCCGATCTAGGTGTAGAGCCACCTTTTCTGATGGTACGAACCTTG 485
Db	75240	TGGCAAGCATGATCCATGAAGTGGGTATTGAAGCGATGTTTCTGATGGGACCAACTCG 75181
Qy	486	TAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAGCGGGCGAAGTGAATTTTG 545
Db	75180	TAACTGTCAATCCCTATTGAGGCTAATGTTAAATGGTTGGTTCTGGTGAATG--TTCT 75124
Qy	546	GTTGCGATAAAGACATCGAGCTCAATGAGGCAAAAGATGAACCGAACTTGAGGTTACTA 605
Db	75123	TAAAAATGAAGACATCACTATCAAGAGGCAAAAAAGCCGTTAGCGTGAAGTTAAAA 75064
Qy	606	ATGAAGGCCCTAAATCCTTGATGGGTAGGCTTCCATTTCTTTGAAAGTAAACAAG 665
Db	75063	ACGTGGGCGACAGACCGGTTCAATTCGGTTTCACTTCCATTTCTTTGAAAGTGAATAGAT 75004
Qy	666	CACATAAATTCGATCGTGAAGAGCCTATGCAAGCGCTAGATATTCCTCTGGCAACA 725
Db	75003	GCCTAGACTTTGACAGAGAAAAAATTTTCGGCAAAACGCTTAGACATTCGAGCGGACAG 74944
Qy	726	CGCTACGCAATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTA 785
Db	74943	CGGTAAAGTTTGAAGCTTGGCGAGAGAAATTCGTTAGATTTGATTGACATTTGTTGGCAACA 74884
Qy	786	AAAAAGTGAATGGCAATGAAACGGCTTTGTGAATPAACATCGGAGTGAACGCCATAACATA 845
Db	74883	GAAGAACTTTTGGATTTAACGCATTTGGTTGATAGGCAAGCAGACAAACGAAGCAAAAAA 74824
Qy	846	AAGCCCTTGACAGCGGAATCTCAGGATTT----- 877
Db	74823	TGTCTTTACACAGACTAAAGAGCGTGGTTTTCATGGTGTAAAAAGCGATGACAACTATG 74764
Qy	878	----ATCAAGTAAGGAGACTCCATGAAA--ATCAAAAAACAAGATATGTAAATACT 930
Db	74763	TAAAAACAAATTAAGAGGTAAAGAAATGAAGAAATTAAGCAAGAAATATGTTCTATGT 74704
Qy	931	ACGGACCCAAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCAAGAGTAG 990
Db	74703	ATGGCCCTACACAGCGGATAAAGTGAATTTGGCGATACAGATTTGATCGTGAAGTAG 74644
Qy	991	AACATGACTATACCACTATGGCGAGAACTTAAATTTGGCGCGGTAAACATCTCCGTG 1050
Db	74643	AACATGACTATACCACTTTATGGCGAAGAGCTTTAAATTCGGTGGCGGTAAAAACCCCTAAGAG 74584
Qy	1051	AGGATATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTTAGTCACTAACG 1110
Db	74583	AAGCATGAGCAATCCAAACCAACCTAGCAAAAGAACTGGATTTGATCATCACTAACG 74524
Qy	1111	CGATGATTATCGACTACACCGGATTTTACAAAGCCGACATTTGGGATTAAAAACGGCAAAA 1170
Db	74523	CTTTAATCGTGGATTACACCGGTATTTTAAAGCGGATTTGGTATTAAAGACGGCAAAA 74464
Qy	1171	TCCATGGCAATTTGGCAAGGAGGAAACAAAGGACATCAAGATGGCGTAAAGCCCTCATATGG 1230
Db	74463	TCGCTGGCAATTTGGTAAAGCGGTAAACAAGACATCAAGATGGCGTTAAAAACAATCTTA 74404
Qy	1231	TCGTGGGTGGGCAACAGACATCTAGCAGGGGAAGTATGATTATTACCGCTGGGGGAA 1290
Db	74403	CGGTGGGTCTGCTACTGAAGCGCTAGCTGGTGAAGGCTTAATCGTAAACGGCTGGTGGTA 74344
Qy	1291	TCGATTTACACACCACTCTCTTCTCCAAACAATTCCTACCGCTCTAGCCCAATGGCG 1350
Db	74343	TTGACACACATCTCACTTTCACCCCAACAATCCCTACAGCTTTTGCNAAGCGGTG 74284
Qy	1351	TTTAAACCATGTTTGGAGCGGCAAGGTCTGTAGATGGCAAGATGGGACTTACTATCA 1410
Db	74283	TAAACAACCATGATTGGTGGCGAAACGGTCTCTGCTGATGGCACTAATGCGACTACTATCA 74224
Qy	1411	CTCCGGGCAATTTGAACTTTGCACCGCATGTTGGCGGCGACAGAGTATTCTATGAATG 1470
Db	74223	CTCCAGGTAGAAGAAACTTTAAATGGATGCTCAGAGCGGCTGAAGAAATATTCTATGAAT 74164
Qy	1471	TGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAAAGTAGAAG 1530
Db	74163	TAGGTTCTTAGCTAAAGGTAACGTTCTTAACGACGAGCTTAGCGGATCAAAATTTGAAG 74104
Qy	1531	CGGCGCGGATTTGGTTTAAATTTGCATGAAGACTGGGGCAACAACCAAGTGGGACTCGATC 1590
Db	74103	CGGTGGCATTTGGCTTTAAATCCACGAGACTGGGGCAACCACTCTCTCTGCAATCAATC 74044
Qy	1591	ACTGCTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGATPACAGTCA 1650
Db	74043	ATGCGTTAGACGTTGCGGCAAAATACGATGTCAAGTCTGATCCATCCACACAGACACTTTGA 73984
Qy	1651	ATGAGCGAGTTTATGTAGATGACCCCTAAATGCAATGAACGGCGCGCATCCATGCGCT 1710
Db	73983	ATGAAGCGGTTTGGTGGAGACACTATGGCAGCTATTGCCGACCGCATATGCACTT 73924
Qy	1711	ACCAATTTGAGGAGCGGTTGAGGACACTCACCTGATTTATCAACCATGGCAGGCGAGC 1770
Db	73923	TCCACACTGAAGCGCTGGTGGCGACACGCTCTGACATTTAAAGTGGCGCGTGAAC 73864
Qy	1771	TCAATATTCTACCTCTCCACACCCCACTATTCCCTATACCATTAATACGGTTGCGAG 1830
Db	73863	ACAACATCTACCGCTTCCACTAACCCCACTATCCCTTTCACTGTGAATACAGAAAGCCG 73804
Qy	1831	HACACTTAGACATGCTCATGACATGCCACCACTAGCAAAACGCATCCGCGAGATTAC 1890
Db	73803	AACACATGACATGCTCATGTTGGTGGCCACTTTGGATAAAAAGCATTAAGAGATGTCC 73744
Qy	1891	AAATTTCTCAAGCGCTATCCCGCCCGCTCTATCCGCGCTGAAGATGTGCTCATGATA 1950
Db	73743	AGTTGCTGATTTCAAGGATCCGCCCTCAAAACCATTTGGGCTGAAGACACTTTGCAATGACA 73684
Qy	1951	TGGGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGTGCAGGCGAAGTGA 2010
Db	73683	TGGGGAATTTCTCAATCACAGTTCTGACTCTCAAGCGATGGGTGCTGGTGGTGAAGTTA 73624
Qy	2011	TTCTCGAACTTTGGCAGACTCGGATTAAGAAATAAAGAAATTTGGTAAAGTCTTCTGAAG 2070
Db	73623	TCACCTAGGACTTTGGCAACAGCTGCAAAAAAATAAAGAAATTTGGCCGCTTGAAGAAAG 73564
Qy	2071	ATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAATACACTATCAACC 2130
Db	73563	AAAAAGCGATTAACGACAACTTCAGGATCAACCGCTCTTGCTAATACACCATTAACC 73504
Qy	2131	CGGCTTTGACCCACCGCGTGAAGAGTATATCGGCTCTGTGGAGAGGGCAAGATCGCCG 2190

Db	73503	CAGCGATCGCTCATGGGATTAGCGAGTATGTAGGTTCTGTAGAAAGTGGCAAAGTGGCTG	73444
Qy	2191	ACTTGGTGGTGGAAATCTCTGCTTTTGGCTGTAACCCAAATCTGATCAAGCGC	2250
Db	73443	ACTTGGTATTGTGGAGTCAGCATTTCTTGGCTGAACCCAAACATGATCAAGCGC	73384
Qy	2251	GTATGGTGGTCTTCTCGAAATGGCGATTCTAAGCGCTGTGCGCACTCCCAACCGG	2310
Db	73383	GATTCATGGCTTAAGCAAAATGGCGATGGAAGCTTCTATCCCTACCCACACCGG	73324
Qy	2311	TTTATTACCGGAATGTTTGGGCATCAGCGAAGCGGAATTTGACACGACATCACTT	2370
Db	73323	TTTATTACAGAGAATGTTGCTCACCATGTAAGCTTAATACATGCAAAACATCACTT	73264
Qy	2371	TTGTTTCAAGTCGCTATGAATGCGTGAAGAAAGCTGGCTTGAAGCGCCGAAAG	2430
Db	73263	TTGTGTCGACGCGCTTATGACAAAGCATTAAGAGAAATTAGGGCTTGAAGACAAG	73204
Qy	2431	TTCTACCGGTCAAAACCTGCGTACATCACTCAACAAAGAGACTTCAAGTTCAACGACAAA	2490
Db	73203	TGTTGCGGTGTAATAATTTGCAAGAACATCACTAAAAGAGACATGCAATTCATGACACTA	73144
Qy	2491	CGGCAAAATCACCGTCGATCCGAAACCTTCGAGGCTTTGTAGTGGCAAACTCTGCA	2550
Db	73143	CGCTCATATGAAGTCAATCTGAACTTACATGCTTCTGCGATGGCAAGAGTCA	73084
Qy	2551	CCTCTAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTCTCTTAGG	2604
Db	73083	CTTCTAACCCAGCTAATAAGTGAGCTGGCTCAACTCTTTAGCATTTTCTTAGG	73030
RESULT 14			
WPCOMMENT			
Sequence split into 17 fragments LOCUS AE000511 Accession AE000511			
Fragment Name Begin End			
AE000511_00 1 110000			
AE000511_01 100001 210000			
AE000511_02 200001 310000			
AE000511_03 300001 410000			
AE000511_04 400001 510000			
AE000511_05 500001 610000			
AE000511_06 600001 710000			
AE000511_07 700001 810000			
AE000511_08 800001 910000			
AE000511_09 900001 1010000			
AE000511_10 1000001 1110000			
AE000511_11 1100001 1210000			
AE000511_12 1200001 1310000			
AE000511_13 1300001 1410000			
AE000511_14 1400001 1510000			
AE000511_15 1500001 1610000			
AE000511_16 1600001 1667867			
LOCUS	AE000511	1667867 bp	DNA circular BCT 27-DEC-2005
DEFINITION	Helicobacter pylori 26695, complete genome.		
ACCESSION	AE000511	AE000523-AE000656	
VERSION	AE000511.1	GI:6626253	
KEYWORDS			
SOURCE	Helicobacter pylori 26695		
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
REFERENCE	1 (bases 1 to 1667867)		
AUTHORS	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.		
TITLE	The complete genome sequence of the gastric pathogen Helicobacter pylori		

JOURNAL	Nature 388 (6642), 539-547 (1997)
PUBMED	9252185
REFERENCE	2 (bases 1 to 1667867)
AUTHORS	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.N., Fujii, C., Bowman, C., Smith, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.
TITLE	Direct Submission
JOURNAL	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REFERENCE	3 (bases 1 to 1667867)
AUTHORS	White, O.
TITLE	Direct Submission
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
COMMENT	On or before Dec 27, 2005 this sequence version replaced gi:2313077, gi:2313090, gi:2313102, gi:2313116, gi:2313131, gi:2313137, gi:2313152, gi:2313162, gi:2313173, gi:2313184, gi:2313196, gi:2313205, gi:2313217, gi:2313230, gi:2313247, gi:2313263, gi:2313275, gi:2313287, gi:2313299, gi:2313310, gi:2313321, gi:2313337, gi:2313349, gi:2313363, gi:2313377, gi:2313391, gi:2313403, gi:2313417, gi:2313430, gi:2313451, gi:2313461, gi:2313475, gi:2313485, gi:2313505, gi:2313518, gi:2313530, gi:2313536, gi:2313554, gi:2313564, gi:2313581, gi:2313593, gi:2313602, gi:2313616, gi:2313628, gi:2313641, gi:2313646, gi:2313663, gi:2313672, gi:2313686, gi:2313703, gi:2313712, gi:2313721, gi:2313730, gi:2313736, gi:2313747, gi:2313759, gi:2313778, gi:2313792, gi:2313802, gi:2313812, gi:2313824, gi:2313834, gi:2313845, gi:2313854, gi:2313869, gi:2313880, gi:2313895, gi:2313907, gi:2313918, gi:2313929, gi:2313944, gi:2313957, gi:2313969, gi:2313982, gi:2314007, gi:2314019, gi:2314028, gi:2314042, gi:2314051, gi:2314060, gi:2314075, gi:2314086, gi:2314103, gi:2314115, gi:2314125, gi:2314139, gi:2314150, gi:2314160, gi:2314173, gi:2314188, gi:2314200, gi:2314216, gi:2314230, gi:2314242, gi:2314256, gi:2314268, gi:2314276, gi:2314293, gi:2314301, gi:2314317, gi:2314327, gi:2314340, gi:2314349, gi:2314360, gi:2314373, gi:2314386, gi:2314400, gi:2314409, gi:2314421, gi:2314436, gi:2314452, gi:2314489, gi:2314502, gi:2314517, gi:2314536, gi:2314547, gi:2314560, gi:2314572, gi:2314581, gi:2314587, gi:2314598, gi:2314609, gi:2314623, gi:2314635, gi:2314645, gi:2314670, gi:2314687, gi:2314700, gi:2314708, gi:2314720, gi:2314733, gi:2314743, gi:2314757, gi:2314771.
FEATURES	Location/Qualifiers
source	1. 1667867
	/organism="Helicobacter pylori 26695"
	/mol_type="genomic DNA"
	/strain="26695"
	/db_xref="taxon:85962"
gene	complement(217..633)
	/locus_tag="HP_0001"
CDS	complement(217..633)
	/locus_tag="HP_0001"
	/note="hypothetical protein; identified by Genemark; putative"
	/codon_start=1
	/transl_table=11
	/product="predicted coding region HP0001"
	/protein_id="AAD07074.1"
	/db_xref="GI:2313078"
	/translation="MATRTQAGAVVELLYAFESNGNEIKTKASSMLEKKIKNNOLA FALSFLNGVLEKINEIDALIEPHLKDWDFKRLGSMEKAILRLGAYEIGTPTQNP111 NECEFLGKLYAEINPTPKFNALIDLSLSKLTQKPLN"
gene	complement(635..1105)
	/locus_tag="HP_0002"
CDS	complement(635..1105)
	/locus_tag="HP_0002"
	/note="similar to GB:L42023 SP:P45149 PID:1007275"

77319 TTGCTTTACACAGAGCTAAAGAGCGTGGTGTTCATGGCGCTAAAAAGCGATGACAACTATG 77260
QY T---ATCAAGTAAGAGAGCTCCCATGAAA---ATCAAAAAACAGAAATATGTAATACCT 930
Db TAAATAAAATTAAGAGAGTAAGAAATGAAAAAGATTAGCAGAAAAAGAAATATGTTCTATGT 77200
QY 931 ACGGACCCACCAAAAGGCGATAAAGTGGCTGTAGGAGATACCGATCTTTGGGCGAGAAGTAG 990
Db 77199 ATGGCCCTACTACAGGCGATTAAGTGAGATTGGGCGATACAGACTTGATCGCTGAAGTAG 77140
QY 991 AACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTG 1050
Db 77139 AACATGACTACCACTATGTCGGAAGAGCTTAAATTCGGTGGCGGTAAAAACCCCTGAGAG 77080
QY 1051 AGGTATGGGTGACAGCAATAGCCCTGATGAAACACCCCTAGATTAGTCTATCACTAAG 1110
Db 77079 AAGGCATGAGCCATTCACAAACACCTTACCAAGAGAAATGCGATTAATCATCACTAAG 77020
QY 1111 CGATGATTATCGACTACACCGGGATTTACAAGCGGACATTTGGGATTAATAACGCGCAAAA 1170
Db 77019 CTTTAATCGTGATTAACCGGTATTTATAAGCGGATATTGGTATTAAAGATGGCAAAA 76960
QY 1171 TCATGGCATTTGGCAAGCGAGAAACAGGACATGCAAGATGGCGTAAAGCCCTCATATGG 1230
Db 76959 TCGCTGCATTTGTAAGCGGTAAACAGACATGCAAGATGGCGTTAAACAACTCTTA 76900
QY 1231 TCGTGGGTGGGCAAGAGCACTAGCAGGGGAAGTATGATTATACCGCTGGGGAA 1290
Db 76899 GCGTAGGTCCTGCTACTGAAAGCCTTAGCCGGTGAAGGTTTGTATGCTAACTGCTGGTGA 76840
QY 1291 TCGATTACACACCCACTTCTTTCTCCAGAAATTTCCCTACCGCTACGCCATGGCG 1350
Db 76839 TTGACACACATCCACTTCATTTTACCCCAACAAATCCCTACAGCTTTTTCAGACGCGTG 76780
QY 1351 TTACAAACATGTTGGAGGGCGCACAGTCTGTAGATGGCAAGATGCGACTACTATCA 1410
Db 76779 TAACAACATGATTGGTGGCGAACTGTCTGTCTGTATGGCACTTAATGCGACTACTATCA 76720
QY 1411 CTCGGGCAATGAACTTGCACGCGATGTTGGCGCAGCAGAGAGTATTCTATGAATG 1470
Db 76719 CTCAGGCAAGAAATTTAAATGGATGCTCAGAGCGGCTGAAGATATTCTATGAAT 76660
QY 1471 TGGCTTTTGGCAAGGCAATAGCTCTAGCAAAACAACTTGTAGAAACAAGTAGAG 1530
Db 76659 TAGGTTTCTGGCTAAAGGTAACGCTTCTACGACGCGAGCTTAGCCGATCAAAATTGAAG 76600
QY 1531 CGGCGGATTTGTTTAAATTTGATCAAGACTGGGGCAACACCAAGTCGATCGATC 1590
Db 76599 CTGGTGGATTGGCTTTTAAATCCAGAACTGGGGCAACACTCTTCTGCAATCAATC 76540
QY 1591 ACTGCTTGAGCGTGCGAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCA 1650
Db 76539 ATCGCTTAGATTGTGACACAAATACGATGTGCAAGTCGCTATCCACACAGACTTTGA 76480
QY 1651 ATGAGCGATTATGATGATGACACCTTAAATGCAATGAACGGCGCGCCATCCATCCCT 1710
Db 76479 ATGAAGCGGTGGTGGGAAGACACTATGGCAGCTATTGGCGGACGCACTATGCACACT 76420
QY 1711 ACCATATTAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGCGAGC 1770
Db 76419 TCCACACTGAAGTGTGGCGGGGACACGCTCTGATATTATTAAGATGAGTGTGTGAAC 76360
QY 1771 TCAATATCTACCTCTCTCCACACCCCACTATTCCCTATACCATTAATACGTTGTCAG 1830
Db 76359 ACAACATTTCTCCGCTTCCACTTAACCCCACTATCCCTTTCATGTGAATACAGAAGCAG 76300
QY 1831 AACACTTAGACATGCTCATGACATGCCACCACTAGCAACAACGCAATCCGGAGGATTAC 1890
Db 76299 AACACATGACATGCTTATGTTGTGCCACCACTTGGATTAAGCAATTAAGAGATGCTC 76240
QY 1891 AATTTCTCAAGCGGTATCCGCCCGGGCTCTATCGCGGCTGAAGATGTGCTCATGATA 1950
Db 76239 AGTTCGTGATTCAAGGATCCGCCCTCAAAACCAATTGGCGCTGAAGACACTTTGTCATGACA 76180

QY 1951 TGGGTGTGATCGGATGACAACTCGGATTCGCAAGCAATGGGCGGTGCGAGGCAAGTGA 2010
Db 76179 TGGGGATTTTCTCAATCACAGCTCTGACTCTCAAGCTATGGGTGCTGGTGAAGTTA 76120
QY 2011 TTCTCTCGAACTTGGGAGACTGGCGATAAGAAATAAAAAAGAAATTTGGTAAAGCTTCTCTGAAG 2070
Db 76119 TCACTAGAACTTGGCAACAGCTGACAAAAAACHAAAAAGAAATTTGGCCGCTTGAAGAAG 76060
QY 2071 ATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCAATAACATATCAACC 2130
Db 76059 AAAAAGGCGATACGCAACTTCAGGATCAACAGCTACTTGTCTAAATACACCATTAACC 76000
QY 2131 CGCTTTGACCCACGCGTGAAGGAGTATATCGGCTCTGTGGAGAGGGCAAGATCGCG 2190
Db 75999 CAGCGATCGCTCATGGGATTAGCGAGTATGTAGGTTCTGTAGAGTGGGCAAGTGGCTG 75940
QY 2191 ACTTGGTGTGGAAATCTGCTCTTTTGGGTAAAAACCCAAAAATCGTCATCAAGGGCG 2250
Db 75939 ACTTGGTATTTGGAGTCCGCAATCTTTGGGTAAAAACCCAAACATGATCATCAAGGGCG 75880
QY 2251 GTATGGTGTCTTCTCTGAAATGGGCGATTCTAAACGCGTCTGTGCCACCTCCCAACCGG 2310
Db 75879 GGTTCATTTGGTTGAGTCAATGGGTGACGCGAACGCTTCTATCCCTACCCCAACACG 75820
QY 2311 TTTATTACGGGAATGTTTGGGCTACGCGCAAGCGGAAATTTTGACACAGCATCACTT 2370
Db 75819 TTTATTACAGAGAAATGTTTGGCTCATCATGATGTAAGCCAAATACGATGCAAAACATCACTT 75760
QY 2371 TGTTTTCCAAAGTCCCTTATGAAATGGGTGCAAGAAAGCTGGGCTTAGAGCGGCAAG 2430
Db 75759 TTGTGTCACAGCGCTTATGACAAAGCAATTAAGAAATTTAGGGCTTTGAAGACAAG 75700
QY 2431 TTCTACCGGTCAAAAACCTGCCGTAAACATCAACAGAAAGACTTCAAGTTTCAACCAAAA 2490
Db 75699 TGTTCGCGTAAATAATTTGCAGAAACATCACTTAAAGAGCATGCAATTCACAGACTA 75640
QY 2491 CGGCAAAATCACCGTCGATCGAATAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCA 2550
Db 75639 CGCTCAATTTGAAGTCAATCTGAAACTTACCATGTGTTGGTGGATGGCAAGAAAGTAA 75580
QY 2551 CCTTAACCCACCTCGCAAGTGCCTTAGCCCGGCTACACTTTCTTCTAGG 2604
Db 75579 CTTCTAAACCGCAATAAAGTAGCTTGGCGCAACTCTTTAGCATTTTCTAGG 75526

RESULT 15

HECUREABCD
LOCUS H.pylori urease 5100 bp DNA linear BCT 26-APR-1993
DEFINITION H.pylori urease (ureA, ureB, ureC, ureD) genes, complete cds.
ACCESSION M60398 X57132
VERSION M60398.1 GI:149007
KEYWORDS urease.
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE
1 (bases 1 to 5100)
AUTHORS Labigne, A., Cussac, V. and Courcoux, P.
TITLE Shuttle cloning and nucleotide sequences of Helicobacter pylori
genes responsible for urease activity
J. Bacteriol. 173 (6), 1920-1931 (1991)
JOURNAL 2001995
PUBMED
COMMENT Original source text: H. pylori DNA.
FEATURES
location/Qualifiers
1..5100
/organism="Helicobacter pylori"
/mol_type="genomic DNA"
/db_xref="taxon:210"
510..1847
/gene="ureC"
510..1847
/gene="ureC"

```

/note="putative"
/codon_start=1
/transl_table=1
/protein_id="AAA25018.1"
/db_xref="GI:149008"
/translation="MKLFGTGVGRKAGVKLTPMFVWRLGIAAGLYFKKHSGQTNKLLI
GKDRKSGYMWENALVGAULTSIGNVIQIGMPTPALAFITEDMRGDMGAMISASHNP
FEDNGIKFNFSGSYKSGYKVEEERAEIEIFHDGELLGSLFVGTESMGSAKADIVDIGYI
AHLHSPKHLNQLSLRIVLDTANGAAKYAVPVFSELGADVLVINDPENGNIHQKQA
GALHPNQSQKVRADIGLAFDGDADRLVVDNLGNI VHGDKLGLVGLVYQKSKQA
LSSOAI VATNNLAKLKEYLSODLELKHCAIGDKTVSECMRLKANKFSGEOSGHIIIP
SDYAKTIGJLVCALQVLSLVESKLVSSVSRNPPELYPNLVLNMLVKQKPPLESLLKGY
SNMLKELDGLCEIRHLIRVSGTENKRIILLESAKOEKLLLESKQELKEFFEGHLC"
1841..2239
/gene="urea"
1841..2239
/gene="urea"
/note="putative"
/codon_start=1
/transl_table=1
/protein_id="AAA25019.1"
/db_xref="GI:149009"
/translation="WLKTKKSLVLFVGFPLIFGVDOAIKYAILSGFVYESVLVDIV
LVFNKGVAFSLLSEGLGLKYLQILLGLGIFILMRQRELFKNHAIEFGWVGAGVSN
VLDRFVHGVDVYVYHYGFDLPPLTSLMS"
2659..3375
/gene="urea"
2659..3375
/gene="urea"
/note="putative"
/codon_start=1
/transl_table=1
/protein_id="AAA25020.1"
/db_xref="GI:149010"
/translation="MKLTPKELDKMLHYAGELAKRKEKGIKLVYEAVALISAHIM
BEARGKTAELMQEGLDILKPDVDMVGSAMIHEVGIEAMPFDKVLVTVHTPLEA
NGKLVGSELFLNKEDITINEGKGAIVKVKXNVDGVPQIGSEHFFFEVNRCLDFDREK
TFGRKLDIASGTAVRFEPGEBSVELIDIGNRRIFGFNALVDRQADNESKKIALHRA
KERGFHGKSDDNVYKTIKE"
3379..5088
/note="putative"
/codon_start=1
/transl_table=1
/protein_id="AAA25021.1"
/db_xref="GI:149011"
/translation="MKKLSRKEYVSMYGPITGDKVRLGDTDLIAEVEHDYTIYGEELK
FGGKTLREGMSQSNPSEKDLIITNALIVDTGIYKADI GIGKKGIAIGKGNK
DMQDGVKNKLVSGPATEALAGEILVIAEGDITHIFISQOIPTPASGVVTMIIGG
TGPDADGTNATTIPGRNLKMLRAEYSGYSMNLGFLAKGNSANDSLAQIEAGTIGF
KIHEDWGTTPSAINHADVADKVDQVAIHTDLINEAGCVEDTMAAIAAGTWHITPATE
GAGGAGHPDI LKVAHEINI LPASTNPITPTVNTAEHMDMLVMVCHILDKSI KEDYQF
ADSRIRPQTIAEDTLHDMGIFPSTSSDSQAMGRVGEVITRTWTQADKNKEPGRLKE
EKGNDNFRIKRYSKYITINPAIAGHSYEVYSGVEVKADVLVWSPAPFGVPKNMII
EKGFTALSQMDANASITPPQPVYIREMFAHGKAKYDANITTFVQAAYDKGIGKEELG
LERQVLPVKNCRNITKDMQFNDDTTAHI EVNPETYHFVVDGKEVTSKPANKVSLAQLF
SIF"

RIGIN

Query Match 35.6%; Score 1024.4; DB 15; Length 5100;
Best Local Similarity 65.3%; Pred. No. 3e-308;
Matches 1682; Conservative 1; Mismatches 852; Indels 39; Gaps 10;

y 67 TAATRCATATTAATTTTAAATAATTAATCTT-ATTATCATATATAATAATATTATTACT 125
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
b 2519 TAACACTAATTCATTTTAAATAATAATTAATAGTTAATGAACGCTTCGTTAATCTTAGTAAA 2578

126 TATATTAAAAAGTTAATAAAAAAGTAAACGAAATTAGGACTATAATCCCATTTGCCTTTAAAA 185
2579 TCAAAACATTGCTACATACATCCACCTTGATTGCGTTATGTTCTTCAAGGAAAAACAC 2638

186 TTATAACAAGGAGTAATAGGTGAACACTCACACCCAAAGACGAAGAAAAGTCTTGTATT 245
2639 TTTAAGAAATAGCAGTAATGATGAAGTAACCTCACCCAAAAGAGTTAGATAAGTTGATGCTCC 2698

```

ORIGIN

Query Match 35.6%; Score 1024.4; DB 15; Length 5100;
Best Local Similarity 65.3%; Pred. No. 3e-308;
Matches 1682; Conservative 1; Mismatches 852; Indels 39;

67	QY	TAATRC	TATATTAA	TTTTTAA	TAAATTA	CTT-ATT	ATCATAT	TATAT	TAATAT	TATTTACT	125
		:	:	:	:	:	:	:	:	:	
2519	Db	TAACTA	CTTAAT	CTTTTAA	ATAATA	TAAATTA	GTAGTAA	TGAAC	GCCTT	CTGTTAAT	2578
126	QY	TATAT	TAAAAAG	TATAAAA	AGTAAG	AAATTA	GAGCTA	TAAAT	TCCCAT	TGCGCTTTAAAA	185
2579	Db	TCAAA	ACAT	TGCTCA	ATCA	TCCAC	CTTGAT	TGCTT	TATGCTT	CAAGGAAAAACAC	2638
186	QY	TTTAA	CACAG	GAGTAA	TAGGT	GAAC	TACAC	CCCAAG	AGCAAG	AAAGTTCTTGTAT	245
2639	Db	TTTAA	GAAATAG	GAGAA	TGACAT	GAACT	CACCC	CAAGAG	TAGTAA	GTGATGTCC	2698

Qy	246	ATTATGGGGCGAAGTGGCTTAGAAGCGGCAAGCAGAGGCGCTTAAGACTCAACCAACCCG	305
Db	2699	ACTACGCTGGAGAATTGGCTTAAAAAAGCGAAAGAAAGGCACTAAAGCTTAACTTATGTAG	2758
Qy	306	AAGCCATTGCTTACATTTAGTGCCCATATATGACGAAGCGCGCTGGAAAAAAGAACCG	365
Db	2759	AAGCAGTAGCTTTGATTAGTGCCCATATATGGAAGAGCGAGCTGGTAAAGAGACTG	2818
Qy	366	TTGCCAGCTTATGGAAGAGTGCATGCTTTTGTGAAAAAGATGAAGTAAATGCCCCGGG	425
Db	2819	CGGCTGAAATTGATGCAAGAGGCGCACTCTTTTAAAAACCAGATGATGTGATGGATGGCG	2878
Qy	426	TGGGTAAATATGTTTCCGATCTAGTGTGTAAGCCACCTTTCTCTGATGGTACGAACCTTG	485
Db	2879	TGGCAAGCATGATCCTCATGAAGTGGGTATTGAAGCGATGTTTCTCTGATGGGACTAAAC	2938
Qy	486	TAACTGTGAATTGGCCCATCGAACCCAGATGAGCACTTCAAAGCGGCGAAGTGAATAATTG	545
Db	2939	TAAACGTTGCATACCCCTATTGAGGCCAATGGTAAATTAGTTCTCTGGTGAAGTTG---TTCT	2995
Qy	546	GTTTCGATTAAGACATCGAGCTCAATCGACGCCAAAGAAAGTAAACCGAACTTGAAGTTACTA	605
Db	2996	TAAAAATGAAGACATCACTATCAAAGAGCAAAAAAGCCGTTAGCGTGAAGTTAAAA	3055
Qy	606	ATGAAGGCGCTAAATCTCTGCATGTGGGTAGCCATTTCCACTTCTTTGAAGACTAAACAAG	665
Db	3056	ATGTTGGCGACAGACCGGTTCAATCGCTCACACTTCCATTTCTTTGAAGTGAATAGAT	3115
Qy	666	CACTAATAATTGATCGTGAAGAAAGCCCTATGCGCAAAAGCCCTAGATATTCCTCTGGCAACA	725
Db	3116	GCCTAGACTTTGACAGAGAAAAAACTTTCGGTAAACGCTTAGACATTTGCGAGCGGGACAG	3175
Qy	726	CGCTACGCATTTGGGGCAGGCAAAACCGCAAGTGCAAGTTGATTCCTTTGGTGGCAGTA	785
Db	3176	CGGTAGATTGTAGCTTGGCGAAGAAAAATCCGTAGAAATTGATTGACATTTGGCGGTAAACA	3235
Qy	786	AAAAAGTGATTGGCATGAACGGGCTTGTGTAATAACATCGCGG---ATGAACGCCA-----	837
Db	3236	GAAGATCTTTTGGATTTAACGCATTTGTTGATAGACAAGCAGACAACGAAAGCAAAAAAA	3295
Qy	838	-----TAAACA-----TAAAGCGC-----TTGACMAAGCG-----AA-----ATCTCAGCGATT	876
Db	3296	TTGCTTTTACACAGAGCTTAAAGAGCGTGGTTTTCATGCGCGCTTAAAGCGATGACAACTATG	3355
Qy	877	T---ATCAAGTATAGGAGACTCCCATGAAA---ATGAAAAACAGATATGTAATAACCT	930
Db	3356	TAAAAACAATTAGGAGTAAGAATGAAAAAGATTAGCAGAAAAGATAATGTTTCTATGT	3415
Qy	931	ACGACCCCAAGGCGATAAAGTGGCTTAGAGATACCGATCTTTGGGCGAAGTAG	990
Db	3416	ATGGTCTTACTACAGGCGATTAAGTGAGATTGGGCGATACAGACTTGATCGCTGAAGTAG	3475
Qy	991	AACATGACTATACCACTATGGGGAAGAACTTAAATTTGGCGGGGTAAACCTATCCGCTG	1050
Db	3476	AACATGACTACACCATTTATGGCGAAGAGCTTAAATTCGCTGGCGGTAAAAACCCCTAAGAG	3535
Qy	1051	AGGATGGGTACAGACAATAGCCCTGATGAAACACCCCTAGATTTTGTATTTAGTCACTAACG	1110
Db	3536	AAGGCATGAGCCCAATCTTAAACACCCCTAGCAAAAGAGTTGGATTTTAAATTATCACTAACG	3595
Qy	1111	CGATGATTATCGACTACACCCGGGATTTTACAAAGCCGACATTTGGGATTTAAAAAAGCGCAAAA	1170
Db	3596	CTTTAATCGTGGATTACCCGGTATTTTAAAGCGGATATTTGGTATTTAAGATGAGCAAAA	3655
Qy	1171	TCCATGCAATTGGCAAGCGAGGAAACAAGGACATGCAAGATGGCGGTAAAGCCCTCATATGG	1230
Db	3656	TCGCTGSCATTTGGTAAAGGCGGTAAACAAGACATGCAAGATGGCGTTTAAAAAACAATCTTA	3715
Qy	1231	TCGTGGGTGTGGGCACAGAGACCTAGCAGGGGAAGGTATGATTATTTACCGCTGGGGGAA	1290
Db	3716	GCGTAGGCTCTGCTACTCAAGACCTTAGCCGCTGAAGGTTTGCATCGTAACGCGCTGGTGTA	3775

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:12:00 ; Search time 1748 Seconds
(without alignments)
11499.424 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2880.6
Sequence: 1 rgragatttccarctt.....aaaaaagtagagccacagg 2883

Scoring table: IDENTITY_NUC
Gapop 3.0 , Gapext 3.0

Searched: 524920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_8:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*
15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2880.6	100.0	2883	7 ADJ58237	Adj58237 Urease su
2	2228	77.3	2452	7 ADJ58249	Adj58249 UreaseXY
3	2158.6	74.9	2405	7 ADJ58240	Adj58240 UreaseXY
4	2158.6	74.9	2407	7 ADJ58246	Adj58246 UreaseXY
5	1948	67.6	2183	7 ADJ58243	Adj58243 UreaseXY
6	1134.4	39.4	8407	12 ADQ37847	Adq37847 H. bizzoz
7	1021.4	35.5	2735	2 AAT44351	Aat44351 H. pylori
8	1008.8	35.0	2619	2 AAQ90180	Aaq90180 Helicobac
9	1008.8	35.0	2619	2 AAT45680	Aat45680 H. felis
10	1007.2	35.0	2619	2 AAQ75319	Aaq75319 Urease ur
11	1005.8	34.9	4824	5 AEB55130	Aeb55130 Salmoneil
12	1005.8	34.9	4824	5 AEB55112	Aeb55112 Salmoneil
13	1004.8	34.9	2767	2 AAO12485	Aao12485 DNA encod
14	940.8	32.7	1710	12 ADQ37850	Adq37850 H. bizzoz
15	902.4	31.3	1719	8 ABA00816	Aba00816 H. felis
16	885.8	30.8	1815	2 AAX14230	Aax14230 H. pylori
17	884.8	30.7	1710	8 ACA34557	Aca34557 Prokaryot
18	881.8	30.6	1707	13 ADU05358	Adu05358 DNA encod

19	881.8	30.6	1707	13	ADU05359	Adu05359 DNA encod
20	880	30.5	1710	9	AAL60579	Aal60579 Helicobac
21	878.4	30.5	1710	2	AAQ04329	Aaq04329 Probe for
22	876.8	30.4	1717	8	ABA00817	Aba00817 H. pylori
23	860.6	29.9	2385	2	AAV62460	Aav62460 Helicobac
24	665.4	23.1	110000	2	AAT42063_05	Continuation (6 of
25	653.6	22.7	1698	8	ACA45232_	ACA45232 Prokaryot
26	650.4	22.6	2341	8	ACA45080	ACA45080 Prokaryot
27	648.2	22.5	1710	10	ADF01572	Adf01572 Bacterial
28	636	22.1	2439	8	ACA46520	ACA46520 Prokaryot
29	624	21.7	3919	13	ADT05424	Adt05424 Haemophil
30	624	21.7	85814	13	ADT05644	Adt05644 Haemophil
31	617.2	21.4	2400	2	AAV19002	Aav19002 Klebsiell
32	617.2	21.4	4768	3	AAZ92097	Aaz92097 Klebsiell
33	612.8	21.3	1878	11	ACH96250	Ach96250 Klebsiell
34	612.2	21.3	1701	8	ACA35953	Aca35953 Prokaryot
35	612.2	21.3	1719	8	ACA34156	Aca34156 Prokaryot
36	604.6	21.0	1704	8	ACA25712	Aca25712 Prokaryot
37	599	20.8	1719	8	ACA53973	ACA53973 Prokaryot
38	598.4	20.8	1704	8	ACA23584	ACA23584 Prokaryot
39	586.4	20.4	1698	8	ACA20910	ACA20910 Prokaryot
40	585.6	20.3	1704	8	ACA26184	Aca26184 Prokaryot
41	582	20.2	1701	8	ACA44198	Aca44198 Prokaryot
42	580.2	20.1	110000	10	ACF67367_09	Continuation (10 o
43	580.2	20.1	110000	10	ACF65384_3	Continuation (4 of
44	580	20.1	1716	10	ACF68224_	ACf68224 Photorhab
45	573.4	19.9	1704	9	ADA30499	Ada30499 DNA encod

ALIGNMENTS

RESULT 1
ADJ58237
ID ADJ58237 standard; DNA; 2883 BP.
XX
AC ADJ58237;
XX
DT 06-MAY-2004 (first entry)
XX
DE Urease subunit polypeptide complex encoding sequence.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
FH Key Location/Qualifiers
FT CDS 206..886
FT /*tag= a
FT /product= "urease polypeptide complex"
FT CDS 897..2603
FT /*tag= b
FT /product= "urease polypeptide complex"
XX

EP1176192-A2.

30-JAN-2002.

11-JUL-2001; 2001EP-00202666.

17-JUL-2000; 2000EP-00202565.

(ALKU) AKZO NOBEL NV.

Kustérs JG, Cattoli G;

WPI; 2002-124384/17.

P-PSDB; ADJ58238, ADJ58239.

Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.

PS	Claim 1; SEQ ID NO 1; 76pp; English.	
XX		
CC	The present invention relates to a novel <i>Helicobacter felis</i> urease X and	
CC	Y subunit polypeptides and immunogenic fragments. The polypeptides are	
CC	used in the manufacture of vaccines against <i>Helicobacter felis</i> infections	
CC	and in diagnostic tests to detect antibodies against <i>Helicobacter felis</i> .	
CC	<i>Helicobacter felis</i> is difficult to grow so it is more convenient to use	
CC	the expression products of the genes encoding the urease X and Y subunits	
CC	in the manufacture of vaccines. The present sequence represents an urease	
CC	subunit polypeptide complex encoding sequence of the invention.	
XX		
SQ	Sequence 2883 BP; 871 A; 642 C; 666 G; 698 T; 0 U; 6 Other;	
	Query Match 100.0%; Score 2880.6; DB 7; Length 2883;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 RGGAGATTTTCCARCACTTCAAGCACATATTGATCTGTGTGGGTGTAATTCR	60
DB	1 RGGAGATTTTCCARCACTTCAAGCACATATTGATCTGTGTGGGTGTAATTCR	60
QY	61 ACTTGTTAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTA	120
DB	61 ACTTGTTAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTA	120
QY	121 TTACTTATATTAATAAGTTAATAAAGTAAGAAATTAGGACTATAATCCATTGCCTT	180
DB	121 TTACTTATATTAATAAGTTAATAAAGTAAGAAATTAGGACTATAATCCATTGCCTT	180
QY	181 TAAATTTAACACAGAGTAAAGTGAAGTAACTACACCAAGAGCAAGAAAGTTCTT	240
DB	181 TAAATTTAACACAGAGTAAAGTGAAGTAACTACACCAAGAGCAAGAAAGTTCTT	240
QY	241 GTTATATTATCGCGGGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCA	300
DB	241 GTTATATTATCGCGGGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCA	300
QY	301 ACCGGAAGCCATTTGCTTACATTTAGTGGCCCATATTATGGAAGCAGCGCGCTGGAAAAA	360
DB	301 ACCGGAAGCCATTTGCTTACATTTAGTGGCCCATATTATGGAAGCAGCGCGCTGGAAAAA	360
QY	361 AACCGTGGCCAGCTTATGGAAGAGTGCATGACCTTTTGAAGAAAGATGAATGCTC	420
DB	361 AACCGTGGCCAGCTTATGGAAGAGTGCATGACCTTTTGAAGAAAGATGAATGCTC	420
QY	421 CGGGGTGGGTAAATGTTTCCGATCTAGGTGTAAGCCACTTTCTGTGATGACGAA	480
DB	421 CGGGGTGGGTAAATGTTTCCGATCTAGGTGTAAGCCACTTTCTGTGATGACGAA	480
QY	481 ACTTGTAACTGTGAATTTGGCCCATCGAACAGATGACACTTCAAGCGGCGAAGTGAA	540
DB	481 ACTTGTAACTGTGAATTTGGCCCATCGAACAGATGACACTTCAAGCGGCGAAGTGAA	540
QY	541 ATTTGGTTGCGATAAAGCATCGAGCTCAATGAGGCAAGAAAGTAACCGAACTTGAGGT	600
DB	541 ATTTGGTTGCGATAAAGCATCGAGCTCAATGAGGCAAGAAAGTAACCGAACTTGAGGT	600
QY	601 TACTAATGAAGGCTTAATCTTTCATGTTGGGTAGCCATTTCCACTTCTTTGAAGCTAA	660
DB	601 TACTAATGAAGGCTTAATCTTTCATGTTGGGTAGCCATTTCCACTTCTTTGAAGCTAA	660
QY	661 CAAGGCATTAATTCGATCGTGAAGAAAGCCCTATGCGAAAGCCCTAGATATTCCTCTGG	720
DB	661 CAAGGCATTAATTCGATCGTGAAGAAAGCCCTATGCGAAAGCCCTAGATATTCCTCTGG	720
QY	721 CAACACGCTACGATTTGGGGCAGGACAAACCCGCAAGTGCAAGTTGCTTTGGTGG	780
DB	721 CAACACGCTACGATTTGGGGCAGGACAAACCCGCAAGTGCAAGTTGCTTTGGTGG	780
QY	781 CAGTAAAAAAGTGATGTCATGAAACCGGCTTGTGAATACATCGCGGATGAACGCCATAA	840
DB	781 CAGTAAAAAAGTGATGTCATGAAACCGGCTTGTGAATACATCGCGGATGAACGCCATAA	840

QY	841 ACATAAAGCGCTTGACAAGGCGAAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGA	900
DB	841 ACATAAAGCGCTTGACAAGGCGAAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGA	900
QY	901 AAATGAAAAAACAAGATATGTAAATACCTACGACCCCAACCAAGGCGATAAAGTGGCT	960
DB	901 AAATGAAAAAACAAGATATGTAAATACCTACGACCCCAACCAAGGCGATAAAGTGGCT	960
QY	961 TAGGAGATACCGATCTTTGGGCGAAGTAGAATACACTATACCTATGCGGAAGAAC	1020
DB	961 TAGGAGATACCGATCTTTGGGCGAAGTAGAATACACTATACCTATGCGGAAGAAC	1020
QY	1021 TTAATTTGGGCGGCTAAACTATCCGTGAGGATGGGTGAGAGCAATAGCCCTGATG	1080
DB	1021 TTAATTTGGGCGGCTAAACTATCCGTGAGGATGGGTGAGAGCAATAGCCCTGATG	1080
QY	1081 AAAACACCTTAGATTTAGTCATCACTAACCGCATGATTAATCGACTACACCGGATTTACA	1140
DB	1081 AAAACACCTTAGATTTAGTCATCACTAACCGCATGATTAATCGACTACACCGGATTTACA	1140
QY	1141 AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCAATTTGGCAAGGCAAGAAC	1200
DB	1141 AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCAATTTGGCAAGGCAAGAAC	1200
QY	1201 ACATCAAGATCGCGTAAGCCCTCATATGCTGCTGGGTGGGCAAGAGCACTAGCAG	1260
DB	1201 ACATCAAGATCGCGTAAGCCCTCATATGCTGCTGGGTGGGCAAGAGCACTAGCAG	1260
QY	1261 GGGAGGTATGATTAATACCGCTGGGGAATCGATTCACACCCACTTCTTTTCTCCAC	1320
DB	1261 GGGAGGTATGATTAATACCGCTGGGGAATCGATTCACACCCACTTCTTTTCTCCAC	1320
QY	1321 AACAAATTCCTACCGCTTAGCCAATGGCGTTTAAACCAATGTTTGGAGCGGCAAGGTC	1380
DB	1321 AACAAATTCCTACCGCTTAGCCAATGGCGTTTAAACCAATGTTTGGAGCGGCAAGGTC	1380
QY	1381 CTGTAGATGGCAAGTAAGTGCATCTATCATCTCGGGGCAAAATGGAATCTGCACCGCATGT	1440
DB	1381 CTGTAGATGGCAAGTAAGTGCATCTATCATCTCGGGGCAAAATGGAATCTGCACCGCATGT	1440
QY	1441 TCGCGGCAAGCAAGATTAATCTATGAATGGGCTTTTGGCAAGGCAATAGCTCTA	1500
DB	1441 TCGCGGCAAGCAAGATTAATCTATGAATGGGCTTTTGGCAAGGCAATAGCTCTA	1500
QY	1501 GCAAAAAACAATCTGTAGAACCAAGTAGAAGCGGCGCATTTGTTTAAATTTGCATGAAG	1560
DB	1501 GCAAAAAACAATCTGTAGAACCAAGTAGAAGCGGCGCATTTGTTTAAATTTGCATGAAG	1560
QY	1561 ACTGGGCGCAACAACAAGTGCATCTGCTTGAAGCGTGGCAGATGAATACGATG	1620
DB	1561 ACTGGGCGCAACAACAAGTGCATCTGCTTGAAGCGTGGCAGATGAATACGATG	1620
QY	1621 TCGAAGTTTGTATCCACACCGGATACGATCAATGAGGCGAGTTATGTAGTACACCCCTAA	1680
DB	1621 TCGAAGTTTGTATCCACACCGGATACGATCAATGAGGCGAGTTATGTAGTACACCCCTAA	1680
QY	1681 ATGCAATGAAGCGGCGCATTCATGCTTACCACTTGAAGGAGCGGGTGGAGGACACT	1740
DB	1681 ATGCAATGAAGCGGCGCATTCATGCTTACCACTTGAAGGAGCGGGTGGAGGACACT	1740
QY	1741 CACCTGATTTATCAACCATGGCAGGCGAGCTCAATTTCTACCCCTCTCCACCAACCCCTAA	1800
DB	1741 CACCTGATTTATCAACCATGGCAGGCGAGCTCAATTTCTACCCCTCTCCACCAACCCCTAA	1800
QY	1801 CTAATTCCTATACCATTAATACGGTTGCAAGACACTTAGACATGCTCATGACATGCCACC	1860
DB	1801 CTAATTCCTATACCATTAATACGGTTGCAAGACACTTAGACATGCTCATGACATGCCACC	1860
QY	1861 ACCTAGACAAACGCAATCCGCGAGGATTTTCAATTTTCTCAAGCGGTATCCGCCCGGCT	1920
DB	1861 ACCTAGACAAACGCAATCCGCGAGGATTTTCAATTTTCTCAAGCGGTATCCGCCCGGCT	1920
QY	1921 CTATCGCGCTGAAGATGTGCTCATGATATGGGTGTGATCGCGATGACAAGCTCGGATT	1980

Db	1921	CTATCGCGCGCTGAAGATGTGCTCATGATATGGGTGTATCGGATGACAAGCTCGGATT	1980
Qy	1981	CGCAAGCAATGGGCGGTGCAGGGGAAGTGAATCTCTCGAACTTGGCAGACTGCGGATAAGA	2040
Db	1981	CGCAAGCAATGGGCGGTGCAGGGGAAGTGAATCTCTCGAACTTGGCAGACTGCGGATAAGA	2040
Qy	2041	ATAAAAAGAAATTTGGTAAAGCTTCTGAAAGTGGCAAGATAAGATAAATTTCCGCATT	2100
Db	2041	ATAAAAAGAAATTTGGTAAAGCTTCTGAAAGTGGCAAGATAAGATAAATTTCCGCATT	2100
Qy	2101	AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTATA	2160
Db	2101	AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTATA	2160
Qy	2161	TCGGCTCTGTGGAAGAGGGCAAGATCCCGACTTTGGTGGTGTGAATCTTCGCCTTTTTTG	2220
Db	2161	TCGGCTCTGTGGAAGAGGGCAAGATCCCGACTTTGGTGGTGTGAATCTTCGCCTTTTTTG	2220
Qy	2221	CGGTAAAAACCAAAATCGTCAATCAAAAGCGCGTATGGTGGTCTTCTCTGAAATGGGCGATT	2280
Db	2221	CGGTAAAAACCAAAATCGTCAATCAAAAGCGCGTATGGTGGTCTTCTCTGAAATGGGCGATT	2280
Qy	2281	CTAACCGGTCTGTGCCACTCCCCCAACCGGTATTATACCGCGAAATGTTTGGGCATCACG	2340
Db	2281	CTAACCGGTCTGTGCCACTCCCCCAACCGGTATTATACCGCGAAATGTTTGGGCATCACG	2340
Qy	2341	GCAAGGCGAAATTTGACACCAAGCATCACTTTTGTGTTTCCAAAGTCGCTATGAAAAATGGCG	2400
Db	2341	GCAAGGCGAAATTTGACACCAAGCATCACTTTTGTGTTTCCAAAGTCGCTATGAAAAATGGCG	2400
Qy	2401	TGAAGAAAAAGCTGGGCTTAGAGCGCCCAAGTCTCTACCGGTCAAAAACTGCGCTAAACATCA	2460
Db	2401	TGAAGAAAAAGCTGGGCTTAGAGCGCCCAAGTCTCTACCGGTCAAAAACTGCGCTAAACATCA	2460
Qy	2461	CCAAGAAAGACTTCAAGTTCAACGACAAAAACGGCAAAAAATCACCGTCGATCCGAAACCT	2520
Db	2461	CCAAGAAAGACTTCAAGTTCAACGACAAAAACGGCAAAAAATCACCGTCGATCCGAAACCT	2520
Qy	2521	TCGAGGTCTTTGTAGATGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTCCCTCTAG	2580
Db	2521	TCGAGGTCTTTGTAGATGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTCCCTCTAG	2580
Qy	2581	CCAGCGCTACACTTTCTTAGGCACAATGCCCCCTTTGGGGCGAGGTATTTTAGGAA	2640
Db	2581	CCAGCGCTACACTTTCTTAGGCACAATGCCCCCTTTGGGGCGAGGTATTTTAGGAA	2640
Qy	2641	TCCTTCATCAACGCACTGCAATCGGTCTTGCGTGGCGATCGGTGCGCTTTAAAAACAAC	2700
Db	2641	TCCTTCATCAACGCACTGCAATCGGTCTTGCGTGGCGATCGGTGCGCTTTAAAAACAAC	2700
Qy	2701	TTTTTCATCTTTAAGCAATCGCCATTTTAAATTAATTTAAATTTCTTAATAATATATAT	2760
Db	2701	TTTTTCATCTTTAAGCAATCGCCATTTTAAATTAATTTAAATTTCTTAATAATATATAT	2760
Qy	2761	TATGCCCTCATTTTTTAAGGAGAAATATGCGTAGGTCTTTGGTATATGCTATGCGGGG	2820
Db	2761	TATGCCCTCATTTTTTAAGGAGAAATATGCGTAGGTCTTTGGTATATGCTATGCGGGG	2820
Qy	2821	TTCTTTTGGTGTGGGCGCAAGGGTATGAAACCCCATCGCTCAAAAAAGTAGAAGCCAC	2880
Db	2821	TTGTTTGGTGTGGGCGCAAGGGTATGAAACCCCATCGCTCAAAAAAGTAGAAGCCAC	2880
Qy	2881	AGG 2883	
Db	2881	AGG 2883	

RESULT 2

RESOLV 2
ADJ58249

ADJ58249 standard; DNA; 2452 BP.

XXII

AC ADJ58249;

XX	06-MAY-2004	(first entry)
DT	UreaseXY subunit encoding sequence #4.	
XX		
DE	immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.	
XX		
KW	Helicobacter felis.	
XX		
OS		
XX	Key	Location/Qualifiers
PH	CDS	48..728
FT		/*tag= a
FT		/product= "urease protein"
FT		739..2445
FT		/*tag= b
FT		/product= "urease protein"
XX		
XX	EP1176192-A2.	
PN		
XX	30-JAN-2002.	
PD		
XX	11-JUL-2001; 2001EP-00202666.	
XX		
PF	17-JUL-2000; 2000EP-00202565.	
XX		
PR	(ALKU) AKZO NOBEL NV.	
XX		
PA	Kusters JG, Cattoli G;	
XX		
PI	WPI; 2002-124384/17.	
XX	P-PSDB; ADJ58250, ADJ58251.	
DR		
DR	Novel Helicobacter felis urease X and Y subunit polypeptides, useful in	
PT	the diagnosis of Helicobacter felis infections and in the preparation of	
PT	vaccines.	
XX		
FS	Disclosure; SEQ ID NO 13; 76pp; English.	
XX		
CC	The present invention relates to a novel Helicobacter felis urease X and	
CC	Y subunit polypeptides and immunogenic fragments. The polypeptides are	
CC	used in the manufacture of vaccines against Helicobacter felis infections	
CC	and in diagnostic tests to detect antibodies against Helicobacter felis.	
CC	Helicobacter felis is difficult to grow so it is more convenient to use	
CC	the expression products of the genes encoding the urease X and Y subunits	
CC	in the manufacture of vaccines. The present sequence represents an	
CC	ureaseXY subunit of the invention.	

RESULT 2

RESOLV 2
ADJ58249

ADJ58249 standard; DNA; 2452 BP.

XXII

AC ADJ58249;

QY 459 CCACCTTTCCTGATGGTACGAACTTGTAACTGTGAATTTGGCCCATCGAACACAGATGACC 518
Db 301 CCACCTTCCCAGATGGCCAACTCGTAACCTGTGAATTTGGCCCATCGAACCTGTATGAC 360
QY 519 ACTTCAAAAGCGGCGAAGTGAATTTGGTTGGATAAAAGACATCGAGCTCAATGACGCA 578
Db 361 ACTTTAAGCGGCGTGAAGTGAATTTGGCTGTGATTAAGACATTTGACTCAACGCGAGTA 420
QY 579 AAGAAGTAACCGAATCTGAGGTTACTAATGAAGGCGCTAAATCTTGCATGTGGGTAGCC 638
Db 421 AGGAAGTTACCGAATCTAGAAGTTACTAACGAAGGACCTAAATCTTGCATGTGGGTAGCC 480
QY 639 ATTTCCACTTCTTTGAAGCTAACAGGCACCTAAATTTGCATGTGAAAGGCTATGCA 698
Db 481 ATTTCCACTTCTTTGAAGCCAAACAAAGCATTTGAATTTGCATCGGGAAGGCTATGCA 540
QY 699 AACGCTAGATATTCCTCTGSCAACCGCTACGCAATTTGGGCGACGACAAACCCGCAAG 758
Db 541 AACGCTAGATATTCCTCTGSCAACACACTACGCAATTTGGGCGACGACAAACCCGTAAG 600
QY 759 TGCAGTTGATTCCTCTTGGTGCAGTAAAAAGTGAITGGCATGAACGGGCTTTGTGAATA 818
Db 601 TGCAGTTAATCCCTCTTGGGCTAGTAAAAAGTGAITGGCATGAACGGGCTTTGTGAATA 660
QY 819 ACATGCGGATGAACGCCATAAACATAAAGCGCTTGACAAAGCGGAAATCTCACGGATTTA 878
Db 661 ATATTGCGGAGAACGCCATAAACAAAGCGCTAGACAAAGCAAAATCTCACGGATTTA 720
QY 879 TCAAGTAGGAGACTCCCATGAATGAATAAACAAGATATGTAATACCTACCGACCC 938
Db 721 TCAAGTAGGAGACTCCCATGAATGAATAAACAAGATATGTAATACCTACCGACCC 780
QY 939 ACCAAGCGGATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGAAGTAGAAATGAC 998
Db 781 ACCAAGCGGATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGAAGTAGAAATGAC 840
QY 999 TATACCACTATGGCGAAGAACTTAAATTTGGCGGGGTAAACATATCCGTGAGGGTATG 1058
Db 841 TATACCACTATGGCGAAGAACTCAAATTCGGTGCAGGTAAACATATCCGTGAGGGTATG 900
QY 1059 GGTCAAGCATAGCCCTGATGAACAAACCCCTAGATTTAGTTCATCACTAAACGCGATGAT 1118
Db 901 GGTCAAGCATAGCCCTGATGAACAAACCCCTAGATTTAGTTCATCACTAAACGCGATGAT 960
QY 1119 ATCGACTACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGCAAAATCCATGCG 1178
Db 961 ATTGACTACCGGGATTTACAAAGCCGACATTTGGCATTTAAAAATGGCAAAATCCATGCG 1020
QY 1179 ATTGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGTCGTGGGT 1080
QY 1239 GTGGGCAAGAGCACTAGAGGGGAAGGTATGATTTTACCGCTGGGGGAATCGAATCA 1298
Db 1081 GTGGGCAAGAGCACTAGAGGGGAAGGTATGATTTTACCGCTGGGGGAATCGAATCA 1140
QY 1299 CACACCACTTCTTTCCTCAACAAATTCCTACCGCTTAGCCCAATGGGGTTACAACC 1358
Db 1141 CACACCACTTCTTTCCTCAACAAATTCCTACCGCTTAGCCCAATGGGGTTACAACA 1200
QY 1359 ATGTTTGGAGGCGGACAGGTCCTGTAGATGGCAGGAATGGCACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGAGGCGGACAGGTCCTGTAGATGGCAGGAATGGCACTATCACTCCGGGC 1260
QY 1419 AAATGGAACTTGACCCGATGTGGCGGAGCAGAAAGATATTCATGAATGTGGGCTTT 1478
Db 1261 AAATGGAACTTGACCCGATGTGGCGGAGCAGAAAGATATTCATGAATGTGGGCTTT 1320
QY 1479 TTGGGCAAGGCAATAGCTCTAGCAAAACAACTTTGTAGACAAAGTAGAGCGGGCGG 1538
Db 1321 TTGGGCAAGGCAATAGCTCTAGTAAAAACAACTTTGTAGAAACAAAGTAGAGCGGGCGG 1380
QY 1539 ATTTGGTTTTAAATTTGATGAAGAATGGGGCACAAACCAAGTGGCATCGATCACTGCTTG 1598

Db 1381 ATTTGGTTTTAAATTTGATGAAGACTGGGGCACAACTCCAGTGGCATCGATCACTGCTTG 1440
QY 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACGATACAGTCAATAGGCA 1658
Db 1441 AGCGTAGCAGATGAATACGATGTGCAAGTTTGTATACACCGATACGTCATAGGCA 1500
QY 1659 GGTATGTAGATGAACACCTAAATGCAATGAACGGCGGCCCATCCATGCTACCACTT 1718
Db 1501 GGTATGTAGATGAACACCTAAATGCAATGAACGGCGGCCCATCCATGCTACCACTT 1560
QY 1719 GAGGAGCGGGTGGAGGACACTCACTGATGTTATACCATGGCAGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGGTGGAGGACACTCACTGATGTTATACCATGGCAGCGAGTGAATATT 1620
QY 1779 CTACCTCTCTCCACACCCCACTATTCCCTATACCATTAATACGGTTGCGAACAACCTTA 1838
Db 1621 CTACCTCTCTCCACACCCCACTATTCCCTATACCATTAATACGGTTGCGAACAACCTTA 1680
QY 1839 GACATGCTCATGACATGCCACCACTAGACAAACCGCATCCGCGAGGATTTACAATTTCT 1898
Db 1681 GACATGCTTATGACCTGCGCACCACTAGATAAACGCTCGCGAGGATCTCCAATTTCT 1740
QY 1899 CAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGGTG 1958
Db 1741 CAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATCGGTG 1800
QY 1959 ATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGGTGAGGCGAAGTGTCTCTCGA 2018
Db 1801 ATCGCGATGACAAGTTCGATTCGCAAGCAATGGGCGCGCTGGGGAAGTGTCTCTAGA 1860
QY 2019 ACTTGGCAGACTCGGATTAAGAAATAAAAGAAATTTGGTAAGCTTCTCTGAAGATGCAAA 2078
Db 1861 ACTTGGCAAACTGCAGACAAAGAAATAAAAGAAATTTGGTAAGCTTCTCTGAAGATGCA 1920
QY 2079 GATAACGATAATTTCCGATTAAGCGCTACATCTCAAAATACATATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAACTTCGCGCATCAACGCTATATCTCAAAATACACCAATTAATCCGCTTTG 1980
QY 2139 ACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGATCCGCGACTTGGTG 2198
Db 1981 ACCCATGGGTGAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGATCCGCGACTTGGTG 2040
QY 2199 GTGTGGAATCTCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGCGGTATGGTG 2258
Db 2041 GTGTGGAATCTCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGCGGTATGGTG 2100
QY 2259 GTCTTCTCTGAAATGGCGATTTAAACGCTCTGTGCCCATCTCCCAACCGGTTTATTAC 2318
Db 2101 GTGTCTCTGAAATGGCGATTTCTAATGCTGTGCCCATCTCTCAGCGGTATTATTAC 2160
QY 2319 CGCGAAATGTTTGGGCATCACGGCAAGGCAAAATTTGACACCGATCACTTTTGTTC 2378
Db 2161 CGCGAAATGTTTGGGCATCACGGCAAGGCAAAATTTGACACCGATCACTTTTGTTC 2220
QY 2379 AAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTCTTACCG 2438
Db 2221 AAAGTCGCTATGAAATGGGTGTGAAGAAAGCTAGGTTTAGAGCGCAAGTGTCTCCC 2280
QY 2439 GTCAAAACTGCGGTAAACATCAACGAAGAAGCTTCAAGTTCAACGAACAAACGGCAAAA 2498
Db 2281 GTGAAAACTGCGGTAAACATCAACGAAGAAGCTTCAAGTTCAACGAACAAACGGCAAAA 2340
QY 2499 ATCAGCGGTGATCCGAAACCTTCAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCAGCGGTGATCCGAAACCTTCAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
QY 2559 CCCACTGCAAGTGCCTCTAGCCAGCGCTACACTTTCTCTTAGGCACAAAT 2610
Db 2401 CCCACTGCAAGTGCCTCTAGCCCAACGCTAACCTTTCTTTTAGGCATAAT 2452


```
ADJ58240
ID ADJ58240 standard; DNA; 2405 BP.
XX
AC ADJ58240;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit encoding sequence #1.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
FH Key Location/Qualifiers
FT CDS 1..681
FT FT /tag= a
FT FT /product= "urease protein"
FT FT 692..2398
FT FT /*tag= b
FT FT /product= "urease protein"
XX
XX EPI176192-A2.
XX
XX 30-JAN-2002.
XX
XX 11-JUL-2001; 2001EP-00202666.
XX
XX 17-JUL-2000; 2000EP-00202655.
XX
XX (ALKU ) AKZO NOBEL NV.
XX
XX Kusters JG, Cattoli G;
XX
XX WPI: 2002-124384/17.
XX
XX P-PSDB; ADJ58241, ADJ58242.
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
XX the diagnosis of Helicobacter felis infections and in the preparation of
XX vaccines.
XX
XX Disclosure; SEQ ID NO 4; 76pp; English.
XX
XX The present invention relates to a novel Helicobacter felis urease X and
XX Y subunit polypeptides and immunogenic fragments. The polypeptides are
XX used in the manufacture of vaccines against Helicobacter felis infections
XX and in diagnostic tests to detect antibodies against Helicobacter felis.
XX Helicobacter felis is difficult to grow so it is more convenient to use
XX the expression products of the genes encoding the urease X and Y subunits
XX in the manufacture of vaccines. The present sequence represents an
XX ureaseXY subunit of the invention.
XX
XX Sequence 2405 BP; 715 A; 571 C; 588 G; 531 T; 0 U; 0 Other;
XX
XX Query Match 74.9%; Score 2158.6; DB 7; Length 2405;
XX Best Local Similarity 93.6%; Pred. No. 0;
XX Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
XX
XX 206 GTGAACTCACCACCAAGAGCGGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAGT 265
XX 1 GTGAACTCACCACCAAGAGCGGCTTAAAGCTTCTGTTATATATGCGGCGAAGTGCT 60
XX
XX 266 AGAAGCGCAAGAGAGGCGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAGT 325
XX 61 AGAAGCGCAAGAGAGGCGCTTAAAGCTCAATCAACCCGAAGCATTGCTTACATTAGT 120
XX
XX 326 GCCCATATTATGAGAGAGCGCGCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
XX 121 GCCCATATTATGAGAGAGCGCGCGTGGAAAAAACCCTTGCTGAACCTTATGGAAGAA 180
XX
XX 386 TGCATGACATTTTGGAAAAAGATGAAGTAATGCCGGGTGGGTATATATGTTCCCGAT 445
XX 181 TGTATGCACTTTTGGAAAAAGATGAGGTGATGCCGGTGTGGGGAATATGTTCCCTGAT 240
```



```
Db 1321 AGAAGCGGCGCGATGGCTTTAAATTGCAATGAAGACTGGGGCAACACCAAGTGGCAT 1380
Qy 1586 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGAGCGTAGCAGATGAATACGATGTGCAAGTTTGTATCCACCGATAC 1440
Qy 1646 AGTCAATGAGCGAGTTATGTAGATGACACCTTAATCAATGAACGGGCGCGCATCCA 1705
Db 1441 GGTCAATGAGCGAGTTATGTAGATGACACCTTAATCGATGAACGGGCGCGCATCCA 1500
Qy 1706 TGCCTACCAATTCAGGCGGCGGTGGAGGACACTCACCTGATGTATACCATGGCAGG 1765
Db 1501 TGCCTACCAATTCAGGCGGCGGTGGAGGACACTCACCTGATGTATACCATGGCAGG 1560
Qy 1766 CGAGCTCAATATTTACCTCTCCACACCCCACTATTCCTATACCATTAATACCGT 1825
Db 1561 CGAGCTCAATATTTACCTCTCCACACCCCACTATTCCTATACCATTAATACCGT 1620
Qy 1826 TGCAGAACACTTAGACATGCTCATGACATGCCACCTAGACAAACGCATCCGCGAGGA 1885
Db 1621 TGCAGAACACTTAGACATGCTCATGACATGCCACCTAGACAAACGCATCCGCGAGGA 1680
Qy 1886 TTTACAATTTTCTCAAGCCGTATCCGCGCGGCTCTATCCGCGCTGAAGATGCTCCA 1945
Db 1681 TCTCCAGTTTTCCTCAAGCCGTATCCGCGCGGCTCTATCCGCGCTGAAGATGCTCCA 1740
Qy 1946 TGATATGGTGTGATCGGATGACAAGCTCGATTCGCAAGCAATGGGCGGTGACGGCGA 2005
Db 1741 TGATATGGTGTGATCGGATGACAAGCTCGATTCGCAAGCAATGGGCGGTGACGGCGA 1800
Qy 2006 AGTGATTCCTCGAACTTGGCAGACTGCGGATGAAGATATAAAGAAATTTGGTAAGCTTC 2065
Db 1801 AGTGATTCCTAGAACTTGGCAAACTGCAGACAAAGAAATAAAGAAATTTGGTAAGCTTC 1860
Qy 2066 TGAAGATGGCAAGATTAACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT 2125
Db 1861 TGAAGATGGTGCAGATTAACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT 1920
Qy 2126 CAACCCCGCTTGAACCCAGCGGTGAGCGAGTATATCGGCTCTGGAAGAGGGCAAGAT 2185
Db 1921 TAATCCCGCTTGAACCCAGCGGTGAGCGAGTATATCGGCTCTGGAAGAGGGCAAGAT 1980
Qy 2186 CGCGACTTGTGTGTGGAATCTCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2245
Db 1981 CGCGACTTGTGTGTGGAATCTCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2040
Qy 2246 AGGCGGTATGTGTGCTTCTGAAATGGGCGATTTCAACGCGTCTGTGCCACTCCCA 2305
Db 2041 AGGCGGTATGTGTGCTTCTGAAATGGGCGATTTCAACGCGTCTGTGCCACTCCCA 2100
Qy 2306 ACCGCTTTATACCGGAAATTTTGGGCATCACGCAAGCGAAATTTGACACCGCAT 2365
Db 2101 GCCGCTTTATACCGCGAAATTTTGGGCATCACGCAAGCGAAATTTGACACCGCAT 2160
Qy 2366 CACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAGCTTAGAGCG 2220
Qy 2426 CCAAGTTCTACCGGTCAAAAATCGCGTAAACATCAACCAAGAAAGCTTCAAGTTCAACGA 2485
Db 2221 CAAGGTCTACCGGTGA AAAATCGCGCAACATCACTAAGAAAGACTTCAAAATCAACAA 2280
Qy 2486 CAAAACGGCAAAATCACCGTGCATCCGAAACCTTCGAGTCTTTGTAGATGGCAACT 2545
Db 2281 CAAGCGCGCATATCACTGTGCATCTTAAACCTTCGAGGCTTTGTAGATGGCAACT 2340
Qy 2546 CTGCACCTCTAAACCACTCGCAAGTGCCTCTAGCCGAGCGCTACACTTTCTCTAGGC 2605
Db 2341 CTGCACCTCTAAACCGGCTCTGAAGTGCTCTTAGCCCAACGCTACACTTTCTCTAGGC 2400
Qy 2606 ACAAT 2610
|||||
```

```
Db 2401 ACAAT 2405
RESULT 4
ADJ58246
ID ADJ58246 standard; DNA; 2407 BP.
XX
AC ADJ58246;
XX
DT 06-MAY-2004 (first entry)
XX
UreaseX subunit encoding sequence #3.
XX
immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
Helicobacter felis.
XX
Key Location/Qualifiers
CDS 2..682
FT /*tag= a
FT /product= "urease protein"
FT 693..2399
FT /*tag= b
FT /product= "urease protein"
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
WPI; 2002-124384/17.
DR P-PSDB; ADJ58247, ADJ58248.
XX
Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
Disclosure; SEQ ID NO 10; 76pp; English.
XX
The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
XX
Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an
ureaseX subunit of the invention.
XX
SQ Sequence 2407 BP; 720 A; 574 C; 584 G; 528 T; 0 U; 1 Other;
Query Match 74.9%; Score 2158.6; DB 7; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
Qy 206 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGTTATATATCGGCGCAAGTGGCT 265
Db 2 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGTTATATATCGGCGCAAGTGGCT 61
Qy 266 AGAAGCGCAAGAGCAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 325
Db 62 AGAAGCGCAAGAGCAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 121
Qy 326 GCCCATATTATGAGCAGAGCGCGCTGGAAGAAACCGTTGCCAGCTTATGGAGAG 385
Db 122 GCCCATATTATGAGCAGAGCGCGCTGGAAGAAACCGTTGCCAGCTTATGGAGAG 181
```


Db	2342	CTGCACCTCTAAACCGCCTCTGGAAGTGCCTCTAGCCAGCGGTACACTTCTTCTTAGGC	2401
Qy	2606	ACAATG 2611 	
Db	2402	NCAATG 2407	
RESULT 5			
ADJ58243	ID	ADJ58243 standard; DNA; 2183 BP.	
XX	AC	ADJ58243;	
XX	06-MAY-2004	(first entry)	
DT	DT	DT	
XX	XX	UreaseXY subunit encoding sequence #2.	
DE	DE	immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.	
XX	KW	Helicobacter felis.	
XX	OS	Location/Qualifiers	
XX	XX	3..683	
XX	XX	/*tag= a	
FT	FT	/product= "urease protein"	
FT	FT	694..2181	
FT	FT	/*tag= b	
FT	FT	/product= "urease protein"	
XX	XX	BP1176192-A2.	
XX	XX	30-JAN-2002.	
XX	11-JUL-2001;	2001EP-00202666.	
XX	17-JUL-2000;	2000EP-00202565.	
XX	(ALKU) AKZO NOBEL NV.		
XX	Kusters JG, Cattoli G;		
XX	WPI; 2002-124384/17.		
DR	P-PSDB; ADJ58244, ADJ58245.		
XX	Novel Helicobacter felis urease X and Y subunit polypeptides, useful in		
PT	the diagnosis of Helicobacter felis infections and in the preparation of		
PT	vaccines.		
XX	Disclosure; SEQ ID NO 7; 76pp; English.		
XX	The present invention relates to a novel Helicobacter felis urease X and		
CC	Y subunit polypeptides and immunogenic fragments. The polypeptides are		
CC	used in the manufacture of vaccines against Helicobacter felis infections		
CC	and in diagnostic tests to detect antibodies against Helicobacter felis.		
CC	Helicobacter felis is difficult to grow so it is more convenient to use		
CC	the expression products of the genes encoding the urease X and Y subunits		
CC	in the manufacture of vaccines. The present sequence represents an		
XX	ureaseXY subunit of the invention.		
SQ	Sequence 2183 BP; 638 A; 505 C; 550 G; 490 T; 0 U; 0 Other;		
Query Match 67.6%; Score 1948; DB 7; Length 2183;			
Best Local Similarity 93.3%; Pred. No. 0;			
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;			
Qy	206	GTGAAACTCACACCCCAAGCAAGAAAGTCTTGTATATTATGCGGCGAAGTGGCT 265	
Db	3	GTGAAACTCACACCCCAAGCAAGAAAGTCTTGTATATTATGCGGCGAAGTGGCT 62	
Qy	266	AGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCAACCGAAGCAATTGCTTACATTAGT 325	
Db	63	AGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAATCAACCCGAAGCAATTGCTTACATTAGT 122	

```
Db 1203 CATCACTCCGGGCAAAATGGAATCTTGCACCGCATGTTGGCGCAGCTGAAGATTTCTAT 1262
Qy 1466 GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAACCTGTAGAACAAAGT 1525
Db 1263 GAATGTAGGCTTTTGGGCAAAAGGCAATAGTTCTAGCAAAAAACAACCTGTAGAACAAAGT 1322
Qy 1526 AGAAGCGGCGGATGTTTAAATTCATGAAGACTGGGGCACAACCAAGTGGCAT 1585
Db 1323 AGAAGCGGCGGATGTTTAAATTCATGAAGACTGGGGCACAACCAAGTGGCAT 1382
Qy 1586 CGATCACTGCTTGAGCGGCGAGATGAATAGATGCAAGTTTGTATCCACACCGATAC 1645
Db 1383 CGATCACTGCTTGAGCGGCGAGATGAATAGATGCAAGTTTGTATCCACACCGATAC 1442
Qy 1646 AGTCAATGAGCGAGTTATGTAGATGACACCTTAAATGAATGAACGGCGCGCCATCCA 1705
Db 1443 GGTCAATGAGCGAGTTATGTAGATGACACCTTAAATGAATGAACGGCGCGCCATCCA 1502
Qy 1706 TGCCTACACATTTAGGGAGCGGTGGAGGACACTCACCTGATGTTATCACATGGGAGG 1765
Db 1503 TGCCTACACATTTAGGGAGCGGTGGAGGACACTCACCTGATGTTATCACATGGGAGG 1562
Qy 1766 CGAGCTCAATATCTACCTCTCCACCCACCCCACTATTCCTATACCATTAATACGGT 1825
Db 1563 CGAGCTCAATATCTACCTCTCCACCCACCCCACTATTCCTATACCATTAATACGGT 1622
Qy 1826 TGCAGAACACCTTAGACATGCTCATGACATGCCACCTAGACAAAGCATCTCCGCGAGGA 1885
Db 1623 TGCAGAACACCTTAGACATGCTCATGACCTGCCACCTAGATAGGCGATCCGCGAGGA 1682
Qy 1886 TTTACAAATTTCTCAAGCCGATCTCCGCGGCTCTATCCGCGCTGAAGATGTCTCCA 1945
Db 1683 TTTACAAATTTCTCAAGCCGATCTCCGCGGCTCTATTCGCGCTGAGGATGTCTCCA 1742
Qy 1946 TGATATGGGTGATCGGATGACAGCTCGGATTCGCAAGCATGGGCGTGCAGGGA 2005
Db 1743 TGATATGGGTGATCGGATGACAGCTCGGATTCGCAAGCATGGGCGTGCAGGGA 1802
Qy 2006 AGTGATTTCTCGAATTTGGGACATGCGGATGAAGATAAAGAAATTTGGTAAGCTTCC 2065
Db 1803 AGTGATTTCTCGAATTTGGGACATGCGGATGAAGATAAAGAAATTTGGTAAGCTTCC 1862
Qy 2066 TGAAGATGGCAAGATACGATATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGGTGCAGATAAGCAACATTTCCGATCAAGCTATATCTCCAAATACACCAT 1922
Qy 2126 CAACCCGCTTTGACCAACCGCGTGAAGGATATATCGGCTCTGTGGAAGGCGCAAGAT 2185
Db 1923 TAATCCGCTTTGACCAACCGCGTGAAGGATATATCGGCTCTGTGGAAGGCGCAAGAT 1982
Qy 2186 CGCGACTTGTGTGGTATCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2245
Db 1983 CGCGACTTGTGTGGTATCTGCTTTTGGCGTAAACCCAAATCGTGATCAA 2042
Qy 2246 AGCGGTATGTTGCTTCTCTGAAATGGGCGATTTAAAGCGCTCTGTGCGCACTCCCA 2305
Db 2043 AGTGGAATGTTGCTTCTCTGAAATGGGCGATTTAAAGCGCTCTGTGCGCACTCCCA 2102
Qy 2306 ACCGTTTATACCGCAAAATTTTGGGCAATTCAGGCAAGCGGCAAAATTTGACACCGAT 2365
Db 2103 GCCGTTTATACCGCAAAATTTTGGGCAATTCAGGCAAGCGGCAAAATTTGACACCGAT 2162
Qy 2366 CACTTTTGTTCGAAGTCG 2385
Db 2163 CACTTTTGTCTCAAGCG 2182
```

RESULT 6
ADQ37847
ID ADQ37847 standard; DNA; 8407 BP.
XX
AC ADQ37847;

```
XX 07-OCT-2004 (first entry)
XX H. bizozeronii urease gene cluster, ureABIEFGH.
XX Urease; urease gene cluster; urease structural gene;
XX urease accessory gene; ureABIEFGH; Helicobacter bizozeronii infection;
XX antibacterial; gene; ds.
XX Helicobacter bizozeronii.
XX US2004142343-A1.
XX 22-JUL-2004.
XX 12-AUG-2003; 2003US-00639273.
XX 16-AUG-2002; 2002US-0404337P.
XX (CHAN//) CHANG Y.
XX (SIMP//) SIMPSON K W.
XX (ZHUJ//) ZHU J.
XX Chang Y, Simpson KW, Zhu J;
XX WPI; 2004-533502/51.
XX GENBANK; AF330621.
XX Novel isolated nucleic acid molecule having urease gene cluster, and
XX conferring on Helicobacter bizozeronii ability to produce urease, useful
XX as vaccine for preventing disease in mammals infected by H.bizozeronii.
XX Claim 2; SEQ ID NO 1; 40pp; English.
XX The invention relates to an isolated nucleic acid molecule conferring on
XX Helicobacter bizozeronii an ability to produce urease, where the nucleic
XX acid molecule is a urease gene cluster comprising at least one urease
XX structural gene and at least one urease accessory gene. The nucleic acid
XX molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
XX invention also relates to an isolated protein encoded by the nucleic
XX acid, a vaccine for preventing onset of disease in mammals infected by H.
XX bizozeronii comprising a nucleic acid and a carrier, and an isolated
XX antibody or its binding portion raised against the nucleic acid. The
XX nucleic acids, proteins and antibodies are useful for vaccinating mammals
XX against onset of disease caused by infection of H. bizozeronii, which
XX involves administering the sequences. The sequences are useful for
XX detecting H. bizozeronii in a sample of tissue or body fluids which
XX involves providing a nucleic acid as an antigen, providing an antibody,
XX or providing a nucleotide sequence as a probe in a nucleic acid
XX hybridisation assay, contacting the sample with the antigen or the probe,
XX and detecting any reaction which indicates that H. bizozeronii is
XX present in the sample. This sequence represents the H. bizozeronii
XX urease gene cluster, ureABIEFGH.
XX Sequence 8407 BP; 2264 A; 1937 C; 1969 G; 2237 T; 0 U; 0 Other;
XX Query Match 39.4%; Score 1134.4; DB 12; Length 8407;
XX Best Local Similarity 68.5%; Pred. No. 3.9e-227;
XX Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;
Qy 183 AAATTTAAACAAGAGGATTAAGTGAACCTCACACCAAGAGCAAGAAAGTTCTTGT 242
Db 2319 AAATTTGGTAGAGGATTTAGGATGAATTAACCCCTTAAGAGCTGGACAAGCTCATGT 2378
Qy 243 TATATTATGGCGGCAAGTGGCTAGAAAGCGCAAGAGAGGCTTAAGCTCAACCAAC 302
Db 2379 TGCAATTATGGCGGCAATTTGGCTAAAAACGCAAGCAAAATGCGCTTAAGCTAAATATA 2438
Qy 303 CCGAAGCCATTGCTTACATTTAGTCCCATATTATGACGAGCGCGCTTGAAGAAAAA 362
Db 2439 CTGAGGCAAGTACGCTCATCATGTCCTCATGATGGAAGAGAGCCCGCTGAGGTAAGAAA 2498
Qy 363 CCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAAGAAAAAGATGAAGTAATGCGCG 422
```


QY 987 GTAGACATGACTATACCACTATGCGGCAAGAACTTAATTTGGCGGGTAAACTATC 1046
Db 915 GTAGACATGACTACACATTTATGGCGAAGAGCTTAAATTCGGTGGCGGTAAACCCCTA 974
QY 1047 CGTGAAGGTATGGGTACAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTCACTACT 1106
Db 975 AGAAGGCATGAGCCCAATCTAAACCTAGCAAGAAGAGTTGGATTTAATATCACT 1034
QY 1107 AACCGGATGATATGACTACACCGGGAATTTACAAAGCCGACATTTGGGATTTAAAAACGGC 1166
Db 1035 AACGCTTTAATCGTGAATTAACACCGGTATTTATAAGCCGATATTTGGTATTAAAGATGGC 1094
QY 1167 AAAATCCATGCAATTTGGCAAGCCAGCAACAGACATGCAAGTGGCGTAAAGCCCTCAT 1226
Db 1095 AAAATCGCTGGCATTTGGTAAAGCGGCTTAACAAAGACATGCAAGATGGCGTTAAAAACAAT 1154
QY 1227 ATGTCGTGGGTGGGACAGAACCACTAGCAGGGGAAGGTATGATTTATACCGCTGGG 1286
Db 1155 CTTAGCGTAGGTCCTGCTACTGAAGCCTTAGCCGCTGAAGTTTGATCGTAACGGCTGGT 1214
QY 1287 GGAATCGATTTACACACCCACTTCCTTTCTCCACAACTAATTCCTACCGCTCTAGCCAAAT 1346
Db 1215 GGTATTGACACACATCCACTTCATTTTCAACCCCAACAAATCCCTACAGCTTTTGCAGC 1274
QY 1347 GCGGTTACACCATGTTTGGAGGGCGGACAGGTCTCTGTAGATGGCAAGATGCGCACTACT 1406
Db 1275 GGTGTAAACCAACATGATTTGGTGGTGAACCGGCTCTGCTGTGGCACTTAATGCGCACTACT 1334
QY 1407 ATCACTCCGGGCAATGAACTTGACACCGCATTTGTCGCGGACAGAGAGATTTCTATG 1466
Db 1335 ATCACTCAGGCAAGAAATTTAAATGATGCTCAGAGCGGCTGAAGAATATTTCTATG 1394
QY 1467 AATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACACTTTGTAGAAACAAGTA 1526
Db 1395 AATTTAGTTCTTGGCTAAAGGTAAACGCTTCTACGATGCGAGCTTAGCCGATCAATTT 1454
QY 1527 GAAGCGGCGGATTTGGTTTAAATGCAATGAAGACTGGGGCAACACCAAGTGGCATC 1586
Db 1455 GAAGCGGCGGATTTGGCTTTAAATTCACGAAGACTGGGGCACCACTCTCTTCTCAATC 1514
QY 1587 GATCACTGTTGAGGCTGGCAGATGAATACGATGCAAGTTGTAAGTTGATCCACCGATACA 1646
Db 1515 AATCATGCGTTAGATGTTGCGGACAAATACGATGTGCAAGTGCCTATCCACACAGACT 1574
QY 1647 GTCAATGAGGAGGTTATGTAGATGACACCCCTTAAATGCAATGAACGGCGGCCCATCCAT 1706
Db 1575 TTGAATGAAGCGGTTGTGTAGAGACACTATGCTGCTATTTGCTGGACGCACTATGAC 1634
QY 1707 GCCTACCACTGAGGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCCACACTGAAGGCGCTGGCGGCGGACACGCTCTCTGATATTATTTAAAGTAGCCGT 1694
QY 1767 GAGCTCAATTTCTACCTCTCCACGACCCCACTATTCCTATACCAATTAACGGTT 1826
Db 1695 GAACACACATTTCTCCGCTTTCCACTAACCCCACTCCCTTTCCACCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTAGACATGCTATGATGACATGCAACCACTAGACAAACGATCCGCGAGAT 1886
Db 1755 GCAGAGCACTGACATGCTTATGTTGTGCGCACCACTTGGATTAAGCAATTAAGAGAT 1814
QY 1887 TTACAATTTTCTCAAAGCCGATATCCGCCCCGCTCTATCGCGCTGAAGATGTGCTCCAT 1946
Db 1815 GTTCAGTTTCGCTGATTTCAAGGATCCGCTCAACCAATTTGGCTGAAGACACTTTGAT 1874
QY 1947 GATATGGGTGATCGGATGACAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGGAA 2006
Db 1875 GACATGGGGATTTTCTCAATCACCACTCTGACTCTCAAGCGATGGCGGCTGTGGGTGAA 1934
QY 2007 GTGATTCCTCGAATTTGCGACTCGGATAGATAAAGAAATTTGGTAAGCTTCT 2066
Db 1935 GTTATCACTAGAACTTTGGCAACAGCTGACAAAAACAAGAAATTTGGCGGCTTTGAA 1994
QY 2067 GAAGATGCAAGATTAACGATAATTTCCGCTTTAAGCGCTACATCTCCAAATACACTATC 2126

Db 1995 GAAGAAAAAGGCGATTAACGAACTTCAGGATCAAAACGCTACTTGTCTAAATACACCAATT 2054
QY 2127 AACCCCGCTTTTACCCACCGCGGTAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGGATTAGCGATATGTAGTTTCAGTAGAGTGGGCAAGTG 2114
QY 2187 GCGGATTTGGGTGTGGAATCTCGCTTTTGGGCTTAAACCCCAAAATCGTGATCAAA 2246
Db 2115 GGTGACTTTGGTATTGTGGAGTCCAGCAATTTCTTTGGCGTGAACCCCAACATGATCATCAA 2174
QY 2247 GCGGATTTGGGTCTCTCTGAAATGGCGGATTTCAACGCGCTCTGTGCGCCACTCCCAA 2306
Db 2175 GCGGATTTGCGTTAAGCCAAATGGCGGATGGAACGCTTCTATCCCTACCCCAAA 2234
QY 2307 CCGGTTTATTACCGGAAATGTTGGGCACTCACGCAAGCGGAAATTTGACACCAAGCATC 2366
Db 2235 CCGGTTTATTACAGAGAAATGTTGCTCATCATGTTAAGCTTAATACGATGCAAAATC 2294
QY 2367 ACTTTGTTTCAAAAGTCCGCTATGAAATGCGGTGAAAGAAAGCTGGGCTTTAGAGCGC 2426
Db 2295 ACTTTGTTGCTCAAGCGGCTTATGACAAAGCAATTAAGAAAGAAATTAGGACTTTGAAAGA 2354
QY 2427 CAAGTTCTACCGGTCAAAAACCTGCGGTACATCACCAAGAAAGACTTCAAGTTCAAGGAC 2486
Db 2355 CAAGTTGTTGCGGTAAAAAATTCAGAAATATACCTAAAAAAGACATGCAATTTCAACGAC 2414
QY 2487 AAAACGGCAAAATCACCGTCCGATCCGAAACCTTCGAGGCTTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCATGTTGAAGTCAATCTGNAATCTACCATGTGTTCTGTGATGCAAGAA 2474
QY 2547 TGCACCTTAAACCCACTCGCAAGTGCCTCTAGCCAGCGGTACACTTTTCTTCTAGG 2604
Db 2475 GTAACTTCTAAACCGCAATAAAGTAGAGCTTTGGGCGCACTCTTTAGCATTTTCTAGG 2532

RESULT 8
AAQ90180
ID AAQ90180 standard; DNA; 2619 BP.
XX AC AAQ90180;
XX AC
DT 25-MAR-2003 (revised)
DT 02-NOV-1995 (first entry)
XX Helicobacter felis urease urea/ureB operon.
XX Urease; UreA; UreB; vaccine; Helicobacter infection; UreI;
KW heat shock protein; ss.
XX Helicobacter felis.
XX Key Location/Qualifiers
RBS 31..37
FT /*tag= b
FT /note= "ureA Shine-Dalgarno site"
FT CDS 43..756
FT /*tag= a
FT /EC_number= "3.5.1.5"
FT /note= "UreA"
FT RBS 756..759
FT /*tag= d
FT /note= "ureB Shine-Dalgarno site"
FT CDS 766..2475
FT /*tag= c
FT /EC_number= "3.5.1.5"
FT /note= "UreB"
XX W09514093-A1.
XX 26-MAY-1995.
XX 19-NOV-1993; 93WO-EP003259.

XX 19-NOV-1993; 93WO-EP003259.
XX (INSP) INST PASTEUR.
XX (INRM) INST NAT SANTE & RECH MEDICALE.
XX Labigne A, Suerbaum S, Ferrero R;
XX WPI; 1995-200383/26.
DR P-PSDB; AAR74336, AAR74337.
XX
XX Immunogenic composition against Helicobacter infection - also gene
PT fragment(s) and protein(s) from Helicobacter urease gene cluster and heat
PT shock protein(s).
XX
XX Claim 10; Fig 3; 128pp; English.
XX
XX The sequence encodes urease UreA and UreB proteins, which are components
CC of a novel immunogenic composition capable of inducing protective
CC antibodies against Helicobacter infection. The composition may include
CC the UreA or UreB proteins, a urease- associated heat shock protein
CC (AAR74338-39) or the UreI protein (AAR74340). The composition is used to
CC prepare a vaccine for humans or animals, especially against H. pylori and
CC H. felis. Antibodies against the proteins may be used for treating
CC Helicobacter infection, and primers/probes to the DNA sequence may be
CC used for detection of Helicobacter infection. (Updated on 25-MAR-2003 to
XX correct PN field.)
XX
XX Sequence 2619 BP; 772 A; 569 C; 665 G; 613 T; 0 U; 0 Other;
Query Match 35.0%; Score 1008.8; DB 2; Length 2619;
Best Local Similarity 66.2%; Pred. No. 5.5e-201;
Matches 1634; Conservative 0; Mismatches 792; Indels 42; Gaps 8;
QY 175 TGCCTTTAAATTAACACAGAGGAGTAATAGTGGAACTCACCCCAAGGAGCAAGAAA 234
DB 13 TACCAATAGAAATTC A-ATAAGGAGTTAGATGAAACTAAAGCCTAAAGAACTAGACAA 71
QY 235 GTTCTTGTTATATTATGCGGCGAAGTGGCTAGAAAAGCGCAAGCAGAGGCTTTAAAGCT 294
DB 72 GTTANTGCTCCATTTATGCGGCGAGATTGGCAGAGAACGCTTGGCGGTGTGTGAACT 131
QY 295 CAACCAACCGGAGCCATTGCTTACATTAGTGGCCATATATTAGACGAGCGCCCGTGG 354
DB 132 CAATTACACCGAAGCGGTGCGCTCATTTAGCGGCGTGTGATGAGAAAAGGCGGTGATGG 191
QY 355 AAAAAAACCGTTCGCCAGCTTATGGAAGAGTGATGACACTTTTGA AAAAAGATGAAGT 414
DB 192 TAAATAAAGCGTGGCGGATTTGATGCAAGAGGCGAGGACTTGGCTTAAAAAAGAAAATGT 251
QY 415 AATGCCCGGGTGGTAAATATGTTCCGATCTAGGTGTAGAACCACTTTTCCTGTATGG 474
DB 252 GATGACGGGTAGCAAGCATGATTCATGAGTGGGATTTGAAGCTACTTCCCGCATGG 311
QY 475 TACGAAACTTGTAACTGTGAATTGGCCCATCGAAACAGATGAGCACTTCAAAGCGGGCGA 534
DB 312 AACCAAGCTTGTAACTATCCACACTCCGGTAGAGGATAATGGCAATTTAGCCCCCGGGA 371
QY 535 AGTGAATTTGGTTGCGATAAAGACATCGAGCTCAATGCGAGGCAAGAGTAACCGAACT 594
DB 372 GGT---CTTCTTAAAAATGAGGACATTACTATTAAACGCGGCAAGAAAGCAATTAGCTT 428
QY 595 TGAGGTTACTAATCAAGGCGCTAAATCTTTCGATGTGGGTAGCCATTTCACACTTTTGA 654
DB 429 GAAAGTGA AAAAATAAAGCGCATGCTCTGTGCAAGTGGGATCAATTTCACACTTTTCGA 488
QY 655 AGCTAAACAAGGCACTAAATTCGATTCGTGAAAAAGCCCTATGGCAAAAGCCCTAGATATTCC 714
DB 489 AGTGAATAAGCTCTTGACTTCGATTCGCGCAAAAGCTTTTGCAAAAGCCCTAGACATTGC 548
QY 715 CTCTGGCAACAGCTACGCAATTGGGCGAGGACAAACCCGCAAAAGTGCAAGTTGATTCCTCT 774
DB 549 ATCTGGAACAGCGTGGCTTTGAAACCCCGGGAGGAAAAAAGTGTGGAATCTATTGACAT 608

QY 775 TGTGGCAGTAAAAAAGTGAATGGGATGAACGGGCTTGTGAATAACATCGGGAT---GA 831
DB 609 CGGCGGGAATAAGCGCATCTATGGCTTTAAATTTCTTTGGTGGATCGCAAGCCGATGCCGA 668
QY 832 ACGCCATAAAC-----ATAAAGCGCTTGACAA-----GGCGAAATCTCAACGG 873
DB 669 TGTAAAAAACTCGGCTTAAACCGCGCTAAGAAAAAAGGTTTGGGTCTGTAACTCGGG 728
QY 874 ATTT--ATCAAGTAAGGAGACTC-----CCATGAAATGA---AAAAACAAGA 916
DB 729 TTGTGAAGCGACTAAAGATAAACAATAAGAAAAACCATGAAAAAGATTTTACGAAAGA 788
QY 917 ATATGTAATACTACCGACCCACCAAAGCGGATAAAGTGGCTTAGGAGATACCGATCT 976
DB 789 ATATGTTTCTATGTATGTGTCCTCACTACCGGGGATCGTGTAGACTCGGCGACACTGATTT 848
QY 977 TTGGGCGAAGTAGAACATGACTATACCACTATGCGGAAGAACTTAAATTTTGGCGGG 1036
DB 849 GATCTTAGAAGTGGAGCATGATTTGACCACTTATGTGTGAAGAGATCAAAATTTGGGGCGG 908
QY 1037 TAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTT 1096
DB 909 TAAACTATCCGTGATGGATGAGTCAACCAATAGCCCTAGCTCTTATGAATTAGATTT 968
QY 1097 AGTCATCACTAAACCGGATGATTTATCGACTACACCGGGATTTACAAAGCGGACATTTGGGAT 1156
DB 969 GGTGCTCACTAAACCGCCCTCATTTGTGACTATATCGGGCAATTTACAAAGCGGACATTTGGGAT 1028
QY 1157 TAAAAACGGCAAAATCCATGGCATTTGGCAAGCGCAAGAACAGGACATGCAAGTGGCGT 1216
DB 1029 TAAAGACGCGAAGATTGACAGGCAATTTGGCAAGCGGCAATAAGGACATGCAAGATGGCGT 1088
QY 1217 AAGCCTCATATGCTGCTGGGTGTGGGCACAGACACTAGCAGGCGGAAAGTATGATTAT 1276
DB 1089 AGATAATTAATCTTTCCGTAGTCTGCTACAGAGGCTTTGGCAGCTGAGGCTTGATTT 1148
QY 1277 TACCGCTGGGGAAATCGAATTCACACCCACTTTCCTTTCTCCACAACTTCCTACTCCG 1336
DB 1149 AACCGCTGTGGCATCGATAGCATATTACATTTATCTCTCCCAACAAATCCTACTGC 1208
QY 1337 TCTACCAATGGCGTTACAAACATGTTGGAGCGGCGACAGGTCTCTGTAGATGGCAGAA 1396
DB 1209 TTTTCCAGCGGGTTACAAACATGATTTGGAGGAGGACAGGACCTTGGGATGGCAGAA 1268
QY 1397 TCGCACTACTATCACTCCGGGCAATGGAATTTGCACCCCATGTTGCGGCGCAGCAGAGA 1456
DB 1269 TCGGACCACTCACTCCCGGACGCGTAATCTTAAAAAGTATGTTGCGGTGACGCGGAGA 1328
QY 1457 GTATTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAACTTGT 1516
DB 1329 ATACGCCATGATCTAGGCTTTTGGCTAAGGGGAATGTGCTTACGAACCTCTTTACG 1388
QY 1517 AGAACAAAGTAGAAGCGGCGGCAATTTGTTTAAATTTGCAATGAAGACTTGGGGCACAACCC 1576
DB 1389 CGATCAGATTGAAGCAGCGGCGGATTTGTTTAAATTCACAGAACTTGGGGAAGCAGCAC 1448
QY 1577 AAGTGGGATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCA 1636
DB 1449 TGCAGCTATTACCACTGCTCAATGTGCGCGATGAATACGATGTGCAAGTGGCTATCCA 1508
QY 1637 CACCGATACAGTCAATGAGCGAGGTTATGTAGATGACACCTTAAATGCAATGAACGGCGG 1696
DB 1509 CACCGATACCTTAAAGAGGCGGCTGTGTAGAGACACCTTAGAGCGGATTTGCCGGCG 1568
QY 1697 CGCCATCCATCCCTACCACTTGAAGGAGCGGGTGGAGACACTACCTGATGTTATCAC 1756
DB 1569 CACCATCATACCTTCCACACTGAAGGCGCTGGGGTGGACACGCTCCAGATGTTATCAA 1628
QY 1757 CATGGCAGCGGAGCTCAATATTTCTACCTCTCCACCCACCCCACTATTTCCCTATACCAT 1816
DB 1629 AATGGCAGGGAATTTAAACATTTCTACCGGCTCTACTAACCCGACCAATTTCTTTCCCAA 1688

Db 2469 GTTCTAGG 2476

RESULT 10

AAQ75319

ID AAQ75319 standard; DNA; 2619 BP.

XX AC AAQ75319;

XX DT 16-OCT-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 22-JUN-1995 (first entry)

XX DE Urease ureA/ureB gene.

XX DE Urease; ureA gene; immunogen; vaccine; diagnostic;

XX KW Helicobacter pylori; heat shock protein; HSP; chaperonin; pILL205; ss.

XX OS Helicobacter felis; ATCC 49179.

XX FT Key

XX FT RBS

XX FT CDS

XX FT Location/Qualifiers

XX FT 31..36

XX FT /*tag= a

XX FT 43..756

XX FT /*tag= b

XX FT /label= UreA gene

XX FT /note= "encodes urease A subunit"

XX FT 756..759

XX FT /*tag= c

XX FT 766..2475

XX FT /*tag= d

XX FT /label= UreB gene

XX FT /note= "encodes urease B subunit"

XX DN WO9426901-A1.

XX PD 24-NOV-1994.

XX PP 19-MAY-1994; 94WO-EP001625.

XX PR 19-MAY-1993; 93EP-00401309.

XX PR 19-NOV-1993; 93WO-EP003259.

XX PA (INSP) INST PASTEUR.

XX PA (INRM) INST NAT SANTE & RECH MEDICALE.

XX PI Labigne A, Suerbaum S, Ferrero R, Thiberge J;

XX WPI; 1995-006797/01.

XX P-PSDB; AAR67371.

XX DNA from Helicobacter pylori and Helicobacter felis - used to develop

XX prods. for detection, treatment and prevention of Helicobacter infection.

XX PS Disclosure; Fig 3; 168pp; English.

XX CC Vaccine compositions include the A and B subunits (given in AAR67371) of

XX CC H. felis urease encoded by the ureA/ureB gene (AAQ75319) region of the

XX CC urease gene cluster of pILL205 (NCNM I-1355), as well as the heat shock

XX CC proteins HSPA (AAR67374) and HSPB (AAR67373) encoded by the urease-

XX CC associated HSP gene cluster region (AAQ75321) of pILL689 (NCNM I-1356).

XX CC Recombinant products are expressed in Escherichia coli. (Updated on 25-

XX CC MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise OS

XX CC field)

XX SQ Sequence 2619 BP; 773 A; 570 C; 664 G; 612 T; 0 U; 0 Other;

Query Match 35.0%; Score 1007.2; DB 2; Length 2619;

Best Local Similarity 66.2%; Pred. No. 1.2e-200;

Matches 1633; Conservative 0; Mismatches 793; Indels 42; Gaps 8;

QY 175 TGCCTTTAAATTTAAACAAGGAGTAAATAGTGAAACTCACACCAAGAGCAAGAAA 234

Db 13 TACCAATAGAAATTCa-ATAAGGAGCTTAGGATGAAACTTAACGCCCTAAAGAACTAGACAA 71

QY 235 GTTCTTGTATTATATGCGGCGAAGTAGGTAGAAAGCGCAAGACGAGAGGCTTAAAGCT 294

Db 72 GTTAATGCTCCATTATGCGGCGAGATTGGCAGAAACGCTTTGGCGCGTGGTGAACCT 131

QY 295 CAACCAACCGAAGCCATTGCTTCAATTAGTCCCATATTATGAGCAGAAAGCGCGCGTGG 354

Db 132 CAATTACACCGAAGCGTGGCGCTCATTTAGCGGCGTGTGATGGAAGGCGCGTGTATGG 191

QY 355 AAAAAAACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGT 414

Db 192 TAATAAAGCGTGGCGGATTGATGCAAGAGGAGGAGCTTGGCTTAAAGAAAGAAATGT 251

QY 415 AATGCCCGGGTGGGTAATATATGTTCCCGATCTAGGTGTAGAACCCACCTTTTCTGATGG 474

Db 252 GATGACCGGTAGCAAGCATGATTCATGAAGTGGGATTGAAGCTTAACCTTCCCGATGG 311

QY 475 TACGAAACTTTGTAAGTGTGAATTGGCCCATCGAATGAGATGAGCACTTCAAGAGGGGCGA 534

Db 312 AACCAAGCTTGTAACTATCACAATCCGCTAGAGGATAATGGCAAAATTAGCCCGCGCGA 371

QY 535 AGTGAATTTTGGTTCGGATTAAGACATCGAGCTCAATGCAGGCAAGAAAGATACCGAAT 594

Db 372 GGT----CTTCTTAAAAAATGAGGACATTAATTAACGCCCGCAAGAGCCATTAGCTT 428

QY 595 TGAGTTTACTAATGAAGGCGCTTAAATCTTTCATGTGGTAGCCATTTCACATTTCTTTGA 654

Db 429 GAAAGTGAATAAAGGCGATCGTCTGTCAGGTGGGATCACATTTCCACTTTCGA 488

QY 655 AGCTAACCAAGGCACCTAAATTTGATGCGTGAATAAAGCCTATGGCAAAAGCGCTAGATATCC 714

Db 489 AGTGAATAAGCTCTTGGACTTCGATCGCGCAAAAAGCTTTTGCACACGCGCTAGACATTC 548

QY 715 CTCTGGCAACAGCTACGATTTGGGCGAGGACAAACCCCGCAAGTGCAGTTGATTCCTCT 774

Db 549 ATCTGGAACAGCGGTGGCTTTGAACCCGGGAGGAAAAAAGTGTGGAATCAATGACAT 608

QY 775 TGGTGGCAGTAAAAAAGTGAATGCGATGAACCGGCTTGTGAATAACATCGCGGAT---GA 831

Db 609 CGGCGGGAATAGCGCATCTATGGCTTAAATTTTGTGGATCGCAAGCGCATGCCGA 668

QY 832 ACGCCATAAAC-----ATAAGCGCTTGACAA-----GGCGAAATCTCACGG 873

Db 669 TGGTAAAAAACTCGGCTTAAACCGCTTAAAGAAAAAGGTTTGGGTCTGTAAATCGCGG 728

QY 874 ATTT--ATCAGTAAGGAGACTC-----CCATGAAAAATGA---AAAAACAAGA 916

Db 729 TTGTGAAGCGACTAAAGATAAAACAAATAGGAAAAAACCAATGAAGAAAGATTTACGAAAGA 788

QY 917 ATATGTAATACCTACGGAACCCACCAAGGCGATAAAGTGGCTTAGGAGATACCGATCT 976

Db 789 ATATGTTCTTATGATGTCCCATCTACCGGGATGTGTAGACTCGGCGACTGATTT 848

QY 977 TTGGCAGAGTAGAACATGACTATACCACTATGCGGAAGAACTTAAATTTTGGCGCGG 1036

Db 849 GATCTTAGAAGTGGAGCATGATTGCACCATTTATGTTGAAGAGATCAAAATTTGGGCGCG 908

QY 1037 TAAACTATTCGTAGGGTATGGGTGAGCAATAGCCCTGATGAAGAAACACCTAGATTT 1096

Db 909 TAAACTATTCGTAGGGTATGGGTGAGTCAACCAATAGCCCTAGCTTTATGAATTAGATTT 968

QY 1097 AGTCATCTACTAAGCGGATGATTATCGACTACACCGGATTTTACAAAGCCGACATTTGGAT 1156

Db 969 GGTGCTCTACTAACGCCCTCATTTGTGACTATAGGGCAATTTTCAAGCCGACATTTGGAT 1028

QY 1157 TAAAAACCGCAAAATTCATGGCAATTTGGCAAGGCGAGGAAACAAGGACATGCAAGATGCGT 1216

Db 1029 TAAAGCGCAAGATTGCGAGGATTTGGCAAGGCGAGCAATAGGACATGCAAGATGCGT 1088

QY 1217 AAGCCCTCATATGTCGTGGGTGGGCAAGACACATAGCGGGGAGGATGATAT 1276

Db 1089 AGATAATAATCTTTGCGTAGGTCTCTACAGAGGCTTTGGCAGCTGAGGCTTGTATGT 1148

QY	2357	CACAGCATCATCTTTGTTTCCAAAGTCGCCTATGAAAAATGGCGTGAAGAGAAAGCTGGG	2416
Db	2229	CACCAATATCACTTTTTCGTGTCCTCCCAAGCGGCTTACAGAGCAGGGATCAAAACAAGAAGCTAGG	2288
QY	2417	CTTAGAGCGGCCAGCTTCTACCGGTCAAAAACTGCCGTAAACATCACCAAGAAAGACTTCAA	2476
Db	2289	GCTAGATCGCGCGGCACCGGCACCTGAAAAACTGTGCAATATCACTAAAAGGACCTCAA	2348
QY	2477	GTTCAACGACAAACCGGCAAAAAATCACCGTCGATCCGAAAAACCTTCGAGGTCTTTGTAGA	2536
Db	2349	ATTCAACAGTGTGACCGGCACATATTGATGTCAACCCCTGAACCTATATAAGGTGAAGTGGGA	2408
QY	2537	TGGCAAACTCTGCACCTCTTAAACCCACTCGCAAGTCCCTCTAGCCACGCGCTACACTTT	2596
Db	2409	TGGCAAGAGGTAACTCTTAAAGCAGCAGATGAATTGAGCCTAGCGCAACTTTATAATTT	2468
QY	2597	CTTCTAGG 2604	
Db	2469	GTTCTAGG 2476	
RESULT 11			
AEB55130			
ID	AEB55130 standard; DNA; 4824 BP.		
XX			
AC	AEB55130;		
XX			
DT	22-SEP-2005 (first entry)		
XX			
DE	Salmonella vector pHUR3 expressing helicobacter antigens, SEQ ID 19.		
XX			
KW	Circular; pHUR3; ds; vaccine; helicobacter pylori infection; urease;		
KW	immune stimulation; antibacterial; Immunostimulant; antiulcer;		
KW	cytostatic; antiinflammatory; gastrointestinal-gen.; gastritis;		
KW	peptic ulcer; carcinoma.		
XX			
OS	Helicobacter pylori.		
OS	Salmonella typhimurium.		
OS	Escherichia coli.		
OS	Synthetic.		
XX			
FN	WO200132014-A2.		
XX			
PD	10-MAY-2001.		
XX			
XX			
PF	01-NOV-2000; 2000WO-US030191.		
XX			
PR	01-NOV-1999; 99US-00431705.		
XX			
PA	(ORAV-) ORAVAX INC.		
XX			
PI	Kleanthous H, Londono-Arcila P, Freeman D;		
XX			
DR	WPI; 2001-343379/36.		
DR	P-PSDB; AEB55113, AEB55114, AEB55115, AEB55116, AEB55117, AEB55118,		
DR	AEB55119, AEB55120, AEB55121, AEB55122, AEB55123, AEB55124, AEB55125,		
DR	AEB55126, AEB55127, AEB55128, AEB55129, AEB55131, AEB55132, AEB55133,		
DR	AEB55134, AEB55135, AEB55136, AEB55137, AEB55138, AEB55139, AEB55140,		
DR	AEB55141.		
XX			
PT	Inducing an immune response against Helicobacter in mammals, useful for		
PT	treating Helicobacter induced gastroduodenal diseases.		
XX			
PS	Disclosure; SEQ ID NO 19; 63pp; English.		
XX			
CC	The invention relates to inducing an immune response against Helicobacter		
CC	in a mammal, comprising mucosally administering to the mammal an		
CC	attenuated Salmonella vector containing a nucleic acid molecule encoding		
CC	a Helicobacter antigen, and parentally administering to the mammal a		
CC	Helicobacter antigen. Also included is an attenuated Salmonella vector		
CC	comprising a nucleic acid molecule encoding a Helicobacter antigen. The		
CC	Helicobacter antigen is a urease, a urease subunit, or its immunogenic		
CC	fragment (encoded by the ureA and ureB genes). The mammal is at risk of		

Disclosure: SEO ID NO 19: 63pp: English:

The invention relates to inducing an immune response against *Helicobacter* in a mammal, comprising mucosally administering to the mammal an attenuated *Salmonella* vector containing a nucleic acid molecule encoding a *Helicobacter* antigen, and parenterally administering to the mammal a *Helicobacter* antigen. Also included is an attenuated *Salmonella* vector comprising a nucleic acid molecule encoding a *Helicobacter* antigen. The *Helicobacter* antigen is a urease, a urease subunit, or its immunogenic fragment (encoded by the *ureA* and *ureB* genes). The mammal is at risk of

CC having but does not have Helicobacter infection or has a Helicobacter
 CC infection. The attenuated Salmonella vector further comprises an htrA or
 CC nirB promoter. The vector can be used in inducing an immune response
 CC against Helicobacter in a mammal. The vector can be used to treat
 CC Helicobacter infection. The vector and the method can be used to treat
 CC Helicobacter induced gastroduodenal diseases, including acute, chronic or
 CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
 CC ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was
 CC constructed expressing UreA and UreB and including an E. coli AmpR gene.
 CC The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is
 CC represented by both AB55112 and AB55130, the peptides expressed by
 CC pHUR3 are shown in the sequence listing to be split between AB55112 and
 CC AB55130, yet Figure 4 shows all the peptides being expressed by the one
 CC sequence, therefore all encoded peptides are cross-referenced to both
 CC sequences.

XX Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;

Query Match 34.9%; Score 1005.8; DB 5; Length 4824;
 Best Local Similarity 65.8%; Pred. No. 2.8e-200;
 Matches 1640; Conservative 0; Mismatches 812; Indels 39; Gaps 10;

Qy	150	AACGAAATAGGACTATAATCCCATGCTTTAAATTTAACACAAG-GAGTAATAGGTG	208
Db	29	AAATGAATCTGACGTACACGCAATTTAGATATTAAATCATCCACAGGAGAGATCTCCATG	88
Qy	209	AAACTCACCCAAAGAGCAAGAAAGTTCTTGTTATATTATGCGGCGAAGTGGCTAGA	268
Db	89	AAACTCACCCAAAGAGTTAGATAAGTTGATGCTCCACTACGCTGGAGAAATGGCTAAA	148
Qy	269	AAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAAGCCATGCTTACATTAAGTCC	328
Db	149	AAACGCAAGAAAGGCAATTAAGCTTAATGTATGAGAGCAGTAGCTTTGATTAAGTCC	208
Qy	329	CATATTATGAGAGCGCGCGTGGAAAAACCGTTGCCAGCTTATGGAAGAGTGC	388
Db	209	CATATTATGGAAGAGCGAGAGCTGGTAAAAAGACTGCGGCTGAATTTGATGCAAGAAGG	268
Qy	389	ATGCACATTTTGAATAAGATGAAGTATGCGCGGGTGGGTAAATATGTTCCCGATCTA	448
Db	269	CGCACTCTTTTAAACCGATGATGTGATGCGGTGGCAAGCATGATCATGAGTGTG	328
Qy	449	GGGTAGAACCCACTTTCTGTAGTGTGACGAAACTTGTAACTGTGAAATGGCCCATCGAA	508
Db	329	GGTATTGAAGCGATGTTTCTGTAGTGGGACTAAACTCGTAAACCGTCAACCCCTATTGAG	388
Qy	509	CCAGATGAGCACTTCAAGCGCGGGAAGTGAATTTGGTTGGATTAAGACATCGAGCTC	568
Db	389	GCCAAATGGTAAATTTAGTTTCTGTGGTGGTTG---TTCTTTAAAAAATGAAGACATCACTATC	445
Qy	569	AATCGAGCAAGAGTAACCGAACTTGAGGTGTACTAATGAAGGCGCTAAATCTCTTGCAT	628
Db	446	AACGAAGCAAAAAGCGGTAGCGTGAAGTTTAAATGTTGGCAGACAGCCGGTTCAA	505
Qy	629	GTGGGTAGCCATTTTCCACTTCTTTGAAGCTAACAGGCACTAAAATTCGATCGTGAATAA	688
Db	506	ATCGGCTCACACTTTCCATTTCTTTGAAGTGAATAGATCGCTAGACTTTTGACAGAGAAAA	565
Qy	689	GCCTATGGCAACCGCTAGATATTCCCTCTCGGCAACACGCTACGCATTTGGGCGAGACAA	748
Db	566	ACTTTTCGGTAAACGCTTAGACATTTGCGAGCGGACAGCGGTGAAGATTTGAGCTGCGCAA	625
Qy	749	ACCCGCAAGTGCAATGATTTCTCTTGGTGGCAGTAAAGAAAGTGGATGGAACCGG	808
Db	626	GAAAAATCCGTAGAAATTAATGACATTTGGCGGTAAACAGAGAAATCTTTGGATTTAACGCA	685
Qy	809	CTTGTGTAATACATCGCGG---ATGAACGCCA-----TAAACA-----TAAAGCG	850
Db	686	TTGGTTGATAGCAAGCAGACAAACGAAGCAAAAAATTTGCTTTTACACAGACTAAGAG	745
Qy	851	C-----TTGACAAAGCG-----AA---ATCTCAGGATTT---ATCAAGTAAAGGAGACTCCC	896
Db	746	CGTGGTTTTCATGGCGCTAAAAGCGATGACAACTATGTAAAAACAATTAAGGAGTAAAGAA	805

Qy	897	ATGAAA---ATGAAAAACAAGAAATATGTAAATACCTACGGACCCACCAAGGCGATAAA	933
Db	806	ATGAAAAAGATTAGCAGAAAAAGATATGTTCTATGTATGGTCTCTACTACAGGCGATAAA	865
Qy	954	GTGCGCTTAGGAGATACCGATCTTTGGGCAAGAGTAGACATGACATATACCACTATGCC	1013
Db	866	GTGAGATTGGGCGATACAGACTTGATCGCTGAAGTAGAACATGACATACCACTTTATGGC	925
Qy	1014	GAAGAACTTAAATTTGGGCGCGGTAAAACTATCCGTGAGGGGTATGGGTGAGCAATAGC	1073
Db	926	GAAAGCTTAAATTCGGTGGCGTAAACCCCTAAGAGAAGGCATGAGCCAACTTAACAAC	985
Qy	1074	CCTGATGAAAAACACCCCTAGATTTAGTCATCACTAAACGCGATGATTTATCGACTACACGGG	1133
Db	986	CTAGCAAAAGAGAGCTTGGATTTAATTTATCACTAACCGCTTTAATCGTGGATTAACCGGT	1045
Qy	1134	ATTTACAAAGCCGACATTTGGGATTTAAACCGCAAAATCCATGGCATTTGGCAGGACGGA	1193
Db	1046	ATTTATAAAGCGGATATTGGTATTAAAGATGGCAAAATCGCTGGCAATTTGGTAAAGCGGT	1105
Qy	1194	AACAGGACATGCAAGATGGCGTAAGCCCTCATATGCTCGTGGGTGTGGGCAACAGAGCA	1253
Db	1106	AACAAGACATGCAAGATGGCGTTAAACCAATCTTAGCGTAGGTCTCTACTGAAGCC	1165
Qy	1254	CTAGCAGGGGAAGTATGATTAATTACCGCTGGGGAAATCGATTTACACACCCACTTTCCTT	1313
Db	1166	TTAGCGGTGAAGGTTTGATCGTAACGGCTGGTGGTATTGAACACACATCCACTTCACTT	1225
Qy	1314	TCTCCAAACAAATTCCTTACCGCTCTAGCCATGGGGTTAGAACCATGTTTGGAGGCGGC	1373
Db	1226	TCACCCCAACAAATCCCTACAGCTTTTGCAGCGGTGAACCAACCATGATTTGGTGGGA	1285
Qy	1374	ACAGTCTGTGTAGATGGCAAGATCGCACTACTATCACTCCGGGCAAAATGGAATTTGCAC	1433
Db	1286	ACCGTCTCTGTGTATGGCACTAATCGCACTACTATCACTCCAGGCGAGAGAAATTTAAA	1345
Qy	1434	CGCATGTTGCGCGCAGCAGAGAGTATTCTATGAATTTGGGGCTTTTGGGCAAGCAAT	1493
Db	1346	TGATGCTCAGAGCGCTGAAGAAATTTCTATGAATTTAGTCTTTTGGCTAAAGGTAAC	1405
Qy	1494	AGCTCTAGCAAAAACAATTGTAGAAACAGTGAAGCGGGCGCGATTTGGTTTAAATTG	1553
Db	1406	GCTTTTAAACGATCGAGCTTAGCCGATCAAAATTTGAAGCGGTGCGATTTGGCAATT	1465
Qy	1554	CATGAAGACTGGGCGACAAACCAAGTGCATCGATCACTGTGTGAGCGTGGCAGATGAA	1613
Db	1466	CAGGAAGACTGGGCGACCACTCTCTGCAATCAATCATGCGGTAGATGTTGGCGHACAA	1525
Qy	1614	TACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGTTATGTAGATGAC	1673
Db	1526	TACGATGTGCAAGTTCGCTATCGCCACAGACACTTTTGAATGAAGCGGTTGTGTAGAGAC	1585
Qy	1674	ACCTTAAATGCAATGAACCGGCGCGCATCCATGCGCTACCAATTTGAGGAGCGGTGGA	1733
Db	1586	ACTATGGCTGCTATTGCTGGAAGCACTATGCACACTTTCCACATGAGGCGCTGGCGG	1645
Qy	1734	GGCACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACC	1793
Db	1646	GGACAGCTCTCTGATATTATTAAGTAGCCGCGTGAACCAACATTTCTCCCGTTTCACT	1705
Qy	1794	ACCCCACTATTTCCTTATACCAATTAATACGGTTGCAAGAACACTTAGACATGTCTCATGACA	1853
Db	1706	RAACCCCACTATTCCTTTTACCGTGAATACAGAAGCAGACACATGACATGCTTATGGTG	1765
Qy	1854	TGCCACCACTAGACAAACGCAATCCGCGAGGATTTTCAATTTTCTCAAGCGGTATCCGC	1913
Db	1766	TGCCACCACTTGGATTAAGAGCAATTAAGAAAGATGTTTCACTGCTGATTTCAAGGATCCGC	1825
Qy	1914	CCCGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATCGGATGACAGC	1973
Db	1826	CCTCAAAACCATTTGCGCTGAAGACACTTTTGCATGATCGGGATTTTCTCAATCACCACT	1885

Qy 1974 TCGGATTGCGAAGCAATGGGGCGTGCAGGCGAAGTGATTCCTCGAATCTGGCAGACTGCG 2033
 Db |||||
 Qy 1986 TCTGACTCTCAAGCGATGGCGGTGGTGAAGTATACATAGACTTGGCAACAGCT 1945
 Db |||||
 Qy 2034 GATAAGATAAAAAAGAAATTTGGTAAGCTTCTGGAAGTGGCAAGATAACGATAATTTTC 2093
 Db |||||
 Qy 1946 GACAAAAACAAGAAAGAAATTTGGCGCTTGAAGAAGAAAAAGGCGATAACGACAACTTC 2005
 Db |||||
 Qy 2094 CGCATTAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGC 2153
 Db |||||
 Qy 2006 AGGATCAAAACGCTACTTGTCTCTAAATACACCATTAACCCAGCGATGCTCATGGGATTAGC 2065
 Db |||||
 Qy 2154 GAGTATATCGGCTCTGTGGAAGAGGCGCAAGATCGCGACTTTGGTGTGGAATCTCTGCC 2213
 Db |||||
 Qy 2066 GAGTATGAGTTTCAGTAGAAGTGGGCAAGTGGCTGACTTGGTATTTGGGAGTCCAGCA 2125
 Db |||||
 Qy 2214 TTTTGGCGGTAAACCCAAATGTGATCAAAAGCGGTATGGTGTCTTCTCTGAAATG 2273
 Db |||||
 Qy 2126 TTCTTTGGCGTGAACCAACATGATCATCAAGGCGGATTCATTCGCTTAAGCCAAATG 2185
 Db |||||
 Qy 2274 GCGGATTCTAAGCGTCTGTGCCCACTCCCAACCGGTTTATTACCGGAAATGTTTGGG 2333
 Db |||||
 Qy 2186 GCGGATGCGAAGCTTCTATCCCTACCCCAACACCGGTTTATTACAGAGAATGTTGCT 2245
 Db |||||
 Qy 2334 CATCAGCGCAAGCGAAATTTGACACGAGCATCTTTGTTTCCAAAGTCGCTATGAA 2393
 Db |||||
 Qy 2246 CATCATGGAAGTAAATACGATGCAACATCACTTTTGTGTTCAAGCGGCTTATGAC 2305
 Db |||||
 Qy 2394 AATGGCGTGAAGAAAGCTGGCTGAGCGCCCAAGTTCTACCGGTCAAAAACCTGCGGT 2453
 Db |||||
 Qy 2306 AAGGCATTAAGAAAGATTAAGACTTGAAGACAAGTGTGCGGTAAAAAATTTGAGA 2365
 Db |||||
 Qy 2454 AACATCAACGAAGAAGCTTCAAGTTCAACGACAAAAACGCAAAAATCACCGTGCATCCG 2513
 Db |||||
 Qy 2366 AATATCACTAAAAAAGACATGCAATTCACGACACTACCGCTCACATGGAAGTCAATCCT 2425
 Db |||||
 Qy 2514 AAAACCTTCAGGCTTTGTAGATGCGCAAACTCTGACCTTAAACCACTCCGCAAGTG 2573
 Db |||||
 Qy 2426 GAAACCTTACCATGTGTTGCTGGTGGCAAGAAAGTAACCTTCTAAACCCAGCCAAATAAGTG 2485
 Db |||||
 Qy 2574 CCTCTAGCCGAGCGCTACACTTCTCTAGG 2604
 Db |||||
 Qy 2486 AGCTTGGCGCAACTCTTTAGCAATTTCTAGG 2516
 Db |||||

RESULT 12
 ID AEB55112 standard; DNA; 4824 BP.
 AC AEB55112;
 XX AEB55112;

22-SEP-2005 (first entry)
 DE Salmonella vector pHUR3 expressing helicobacter antigens, SEQ ID 1.
 XX Circular; pHUR3; ds; vaccine; helicobacter pylori infection; urease;
 KW immune stimulation; antibacterial; Immunostimulant; antiulcer;
 KW cytosolic; antiinflammatory; gastrointestinal-gen.; gastritis;
 KW peptic ulcer; carcinoma.

XX Helicobacter pylori.
 OS Salmonella typhimurium.
 OS Escherichia coli.
 OS Synthetic.

XX W0200132014-A2.
 FN 10-MAY-2001.

XX 01-NOV-2000; 2000WO-US030191.
 XX 01-NOV-1999; 99US-00431705.
 XX

PA (ORAV-) ORAVAX INC.
 XX Kleanthous H, Londono-Arcila P, Freeman D;
 PI WPI; 2001-343379/36.
 XX P-PDB; AEB55113, AEB55114, AEB55115, AEB55116, AEB55117, AEB55118,
 DR AEB55119, AEB55120, AEB55121, AEB55122, AEB55123, AEB55124, AEB55125,
 DR AEB55126, AEB55127, AEB55128, AEB55129, AEB55131, AEB55132, AEB55133,
 DR AEB55134, AEB55135, AEB55136, AEB55137, AEB55138, AEB55139, AEB55140,
 DR AEB55141.

Inducing an immune response against Helicobacter in mammals, useful for treating Helicobacter induced gastroduodenal diseases.

Disclosure; SEQ ID NO 1; 63pp; English.

XX The invention relates to inducing an immune response against Helicobacter in a mammal, comprising mucosally administering to the mammal an attenuated Salmonella vector containing a nucleic acid molecule encoding a Helicobacter antigen, and parenterally administering to the mammal a Helicobacter antigen. Also included is an attenuated Salmonella vector comprising a nucleic acid molecule encoding a Helicobacter antigen. The Helicobacter antigen is a urease, a urease subunit, or its immunogenic fragment (encoded by the ureA and ureB genes). The mammal is at risk of having but does not have Helicobacter infection or has a Helicobacter infection. The attenuated Salmonella vector further comprises an htrA or htrB promoter. The vector can be used in inducing an immune response against Helicobacter in a mammal. The vector can be used to treat Helicobacter infection. The vector and the method can be used to treat Helicobacter induced gastroduodenal diseases, including acute, chronic or atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was constructed expressing UreA and UreB and including an E. coli AmpR gene. The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is represented by both AEB55112 and AEB55130, the peptides expressed by pHUR3 are shown in the sequence listing to be split between AEB55112 and AEB55130, yet Figure 4 shows all the peptides being expressed by the one sequence, therefore all encoded peptides are cross-referenced to both sequences.

XX Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;
 Query Match 34.9%; Score 1005.8; DB 5; Length 4824;
 Best Local Similarity 65.8%; Pred. No. 2.8e-200;
 Matches 1640; Conservative 0; Mismatches 812; Indels 39; Gaps 10;

Qy 150 AACGAATTAGGACTATATCCCATGCTTTAAATTTAACACAAG-CAGTAATAGGTG 208
 Db |||||
 Qy 29 AATGAATCTGACGTACACAGCAATTTAGATATTAATCATCCAGGAGAGATCTCCATG 88
 Db |||||
 Qy 209 AAATCTCACACCCAAAGAGCAAGAAAGTTCTTGTATTATATGCGGGCGAAGTGGCTAGA 268
 Db |||||
 Qy 89 AATCTACCCCAAAAGAGTTAGATAGTTGATGCTCCACTAGCTGGAGAAATGGCTAAA 148
 Db |||||
 Qy 269 AAGCGCAAGAGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTTACATTAAGTCC 328
 Db |||||
 Qy 149 AAACGCAAGAAAGGCAATTAAGCTTAATGTATGAGAAGCAGTAGCTTTGATTAGTGC 208
 Db |||||
 Qy 329 CATATTATGACGAAGCGCGCTGGAAAAAACCCTTGCACAGCTTATGGAAGAGTGC 388
 Db |||||
 Qy 209 CATATTATGGAAGAGCGAGAGCTGGTAAAAAGACTGCGGCTGAATTTGATGAAGAAGG 268
 Db |||||
 Qy 389 ATGCACCTTTTGAAGAAAGATGAAGTAATGCCGGGTGGGTAATATGTTCCCGATCTA 448
 Db |||||
 Qy 269 CGCACTCTTTTAAACCCAGATGATGATGATGCGGTGGCGAAGCATGATCCATGAGTG 328
 Db |||||
 Qy 449 GGTGTAGAAGCCACTTTCTCTGATGTACGAAACTTGTAACTGTGAATTTGGCCCATCGAA 508
 Db |||||
 Qy 329 GGTATTGAAGCGATGTTCTCTGATGGGACTAACTCGTAACCGTGCATACCCCTATTAG 388
 Db |||||
 Qy 509 CCAGATGAGCACTTCAAGCGGCGGAAGTGAATTTGGTTGGGATAAGACATCGAGCTC 568
 Db |||||
 Qy 389 GCCAATGTAAATTAGTTTCTCTGGTGGTGTG---TTCTTAAAAAATGAAGACATCACTATC 445
 Db |||||

QY	569	AATG	CAGG	CAAA	GAA	AGTAA	CCGA	ACTTG	AGGTTT	CTAAT	GAA	GGCG	CTAAAT	CTCTT	GCAT	628
Db	446	AACG	AAAG	CAAAA	AAAG	CCGTT	AGCGT	CAAA	AGTTT	AAAAAT	TGTT	GGC	AC	GAC	GACG	505
QY	629	GTGG	TAGC	CAATTT	CCACT	CTCTT	TGA	AGCTT	AA	CAAG	GC	CTA	AAAA	TTCC	GAT	688
Db	506	ATCG	CT	CAC	TTCCA	CTTTCTT	TGA	AGTGA	TAG	ATG	CGCT	AG	CTTT	GAC	AG	565
QY	689	GCCT	ATGG	CAAA	CGCCT	AGATAT	TC	CTCT	GTG	CAAA	CA	CGT	AC	GCT	ATGG	748
Db	566	ACTT	TCG	TAA	CGCT	TAG	ACA	TTG	CG	CGG	CA	CG	CGT	TA	AGTT	625
QY	749	ACCG	CAAA	AGT	GAG	TG	TGAT	TC	CTT	GTG	G	CA	GTAA	AAAA	AGT	808
Db	626	GA	AAAA	TCC	GT	AGAA	TTG	AT	TG	CA	TTGG	CGG	TAA	CAG	AGA	685
QY	809	CTTG	TGA	ATA	CA	T	CGG	---	ATGA	AC	GCA	-----	TAA	CA	-----	850
Db	686	TTGG	TG	TAT	AG	CA	AG	CA	GA	CA	AG	CA	AAAA	AA	TTG	745
QY	851	C----	TTGA	CA	AG	CG	-----	AA	---	AT	CT	C	AG	ATTT	---	896
Db	746	CGTG	TTTT	TC	TG	CG	CT	TAA	AG	CG	AT	G	CA	CT	AT	805
QY	897	ATGA	AAA	---	ATGA	AAAA	CA	AG	AT	AT	GT	AA	TAC	CT	ACG	953
Db	806	ATGA	AA	AG	AT	TAG	CA	AA	AG	AA	AT	AT	GT	TT	CT	865
QY	954	GTG	CG	CT	TAG	GAT	AC	CG	AT	CTTT	TTG	GG	CA	AG	TAG	1013
Db	866	GTG	AT	TG	GG	CA	T	CA	G	AT	CT	G	CT	GA	GT	925
QY	1014	GA	GA	AA	CT	TAA	ATTT	TG	CG	CGG	TAA	AA	CT	AT	CC	1073
Db	926	GA	AG	CT	TAA	TT	CG	TG	CG	G	TAA	AA	CC	T	AA	985
QY	1074	CTG	AT	GA	AA	CA	CC	TAG	AT	T	AG	T	CA	CT	AA	1133
Db	986	CCT	AG	CA	AA	G	AG	TT	G	GA	TTT	AA	TT	T	AC	1045
QY	1134	ATT	T	CA	AA	CG	CA	AT	TG	GA	TTT	AAAA	CG	CA	AA	1193
Db	1046	ATT	T	T	AA	CG	GA	T	TG	GA	TTT	AA	AG	TG	CA	1105
QY	1194	A	CA	AG	G	A	C	A	TG	CG	T	A	AG	C	C	1253
Db	1106	A	CA	AA	G	A	C	A	TG	CG	T	T	AA	AA	CA	1165
QY	1254	CT	AG	C	GG	G	A	AG	T	AT	T	A	C	CG	T	1313
Db	1166	TT	AG	C	CG	T	GA	AG	TT	TG	AT	CG	T	GA	TT	1225
QY	1314	T	C	T	C	A	CA	AA	T	T	C	C	T	AG	C	1373
Db	1226	T	C	A	C	C	C	A	AA	T	T	T	G	C	A	1285
QY	1374	A	C	AG	G	T	C	T	AG	A	TG	CA	T	A	C	1433
Db	1286	A	C	CG	T	C	T	G	A	TG	G	C	A	T	A	1345
QY	1434	C	G	A	TG	T	G	C	G	C	A	G	A	G	A	1493
Db	1346	T	G	A	T	G	T	C	AG	A	A	T	T	C	A	1405
QY	1494	A	G	C	T	AG	C	A	A	A	A	A	A	A	A	1553
Db	1406	G	C	T	T	A	A	C	G	A	T	AG	C	C	A	1465
QY	1554	C	A	T	A	G	A	C	T	G	G	G	C	A	A	1613
Db	1466	C	A	G	A	G	A	C	T	G	G	G	C	A	A	1525

RESULT 13
AAQ12485
ID AAQ1
XX

AAQ12485
ID AAQ12485 standard; DNA: 2767 BP.

XX

1614	QY	TACGATGTCGAAGTTGTTATCCACACCGATACAGTCAATGAGGCAGGTTATGTAGATGAC	1673
1526	DB	TACGATGTCGAAGTCGTATCGCCACAGACACTTTTGAATGAAGCCGGTTGTGTAGAAGAC	1585
1674	QY	ACCTAAATGCAATGAACGGCGCGCCATCCATGCGCTACCACTTGAAGGAGCGGCTGGA	1733
1586	DB	ACTATGGCTGCTATTGCTGGACGGCACTATGCACTTTCCACACTGAAGCGCTGSGCGC	1645
1734	QY	GGACACTCACTGATGTTATCACCATGGCAGGCGAGCTCAATATTCTACCCCTCCTCOACC	1793
1646	DB	GGACACGCTCCTGATATTATAAGTAGCCGCTGAACACAACATTCTTCGCGCTTCCACT	1705
1794	QY	ACCCCACTATTCCCTATACCACTTANTAGCGTTGCGAGAACACTTAGACATGCTCATGACA	1853
1706	DB	AAACCCACCATCCCTTTCACCGTGAATACAGAAGCAGAGACATGAGACATGCTTATGGTG	1765
1854	QY	TGCCACCACTAGACAAACGGCATCCCGGAGGATTTACAATTTTCTCAAGCCGTATCCCGC	1913
1766	DB	TGCCACCACTTGGATAAAGCATTAAGAAGATGTTTCAGTTCCGCTGATTCAAGGATCCGC	1825
1914	QY	CCGGGCTCTATCGCGCTGAAGATGTCCTCAATGATATGGGTGTGATTCGGATGACAAAGC	1973
1826	DB	CCTCAAAACCAATTGGCGCTGAAGACACTTTTGTCATGACATGGGGATTTTCTCAATCACCAGT	1885
1974	QY	TCGGATTCCCAAGCAATGGGCGTGCAGCGAAGTCATTCTCTGAATTTGGCAGACTGCG	2033
1886	DB	CTGACTCTCAAGCGATGGCGCGTGTGGGTGAAGTTATCACTAAGACTTTGGCAACAGCT	1945
2034	QY	GATAAGAATAAAAAAAATTTGGTAAAGTCTCTCTGAAGATGGCAAGATAACGATAAATTC	2093
1946	DB	GACAAAAACAAGAAAGAAATTTGGCGCTTTGAAAGAAAGAAAGCGGATTAACGACAATTC	2005
2094	QY	CGCATTAAGCGCTACATCTCCAAATAACATATCAACCCCGCTTTGACCCACGCGGTGAGC	2153
2006	DB	AGGATCAAAACGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAGC	2065
2154	QY	GAGTATATCGGCTCTGTGGAAGAGGCGCAAGATCGCGCACTTTGGTGTGTGGAATCTCTGCC	2213
2066	DB	GAGTATGTAGTTTCAGTAGAAGTGGGCAAGTGGCTGACTTTGGTATTTGGAGTCCAGCA	2125
2214	QY	TTTTTTTTGGGTTAAACCCCAAAATCGTGATCAAAAGCGGTATGGTGGTCTTCTCTGAAATG	2273
2126	DB	TTCTTTTGGCTGAAACCCCAACATGATCATCAAAAGCGGATTCATTGCGTTAAGCCAAATG	2185
2274	QY	GGCGATTCTAAACGCTCTGTGCCACTCCCAACCGGTTTATTACCGCGAATAATGTTTGGG	2333
2186	DB	GGCGATGCGAAGCTTCTATCCCTACCCCAACAACCGGTTTATTACAGAGAAATGTTTCGCT	2245
2334	QY	CATCAGCGCAAGCGCAAAATTTTGACACGACATCACTTTTGTGTTTCCAAAGTCGCTATGAA	2393
2246	DB	CATCATGTTAAGCTAAATACGATGCAACATCATCTTTTGTGTCTCAGCGGCTTATGAC	2305
2394	QY	AATGGCGTGAAGAAAGCTGGGCTTTAGACGCGCAAGTTCTTACCGGTCAAAAACTGCGGT	2453
2306	DB	AAAGGCATTAAGAAAGAAATTAGGACTTTGAAGACAAGTGTGCGCGTTAAAAAATTCGAGA	2365
2454	QY	AACATCACAAGNAAGACTTCAGTTTCAGGACAAAACGGCAAAATCAACGTCGATCCG	2513
2366	DB	AATATCACTAAAAAAGACATGCAATTTCAACGACACTACCGCTCACATTTGAAGTCAATCCT	2425
2514	QY	AAAACTTCGAGTCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCACTCGCAAGTG	2573
2426	DB	GAAACTTACCATGTGTTGTGGATGGCAAGAAGTAACCTCTTAAACCGACCAATAAAGTG	2485
2574	QY	CTCTTAGCCACGCTACACTTTCTTCTAGG	2604
2486	DB	AGCTTGGCGCACTCTTTAGCATTTTCTTAGG	2516

AC AAQ12485;
 XX
 DT 27-AUG-2003 (revised)
 DT 23-SEP-1991 (first entry)
 XX
 XX DNA encoding A and B subunits of *H. pylori* urease.
 XX
 KW Gastritis; peptic ulceration; duodenitis; helicobacter; campylobacter; ss.
 XX
 OS Helicobacter pylori.
 XX
 FH Key Location/Qualifiers
 FT CDS 64..717
 FT /*tag= a
 FT /label= subunit A
 FT /note= "66 kD"
 FT CDS 721..2400
 FT /*tag= b
 FT /label= subunit B
 FT /note= "31 kD"
 XX
 PN WO9109049-A.
 XX
 XX 27-JUN-1991.
 PD
 XX 19-DEC-1989; 89GB-00028625.
 PF
 XX 19-DEC-1989; 89GB-00028625.
 PR
 XX (THRE-) 3I RES EXPL LTD.
 XX
 PA Tabaqchal IS, Clayton CL, Wren BW, Kleanthous H;
 PI
 XX WPI; 1991-208084/28.
 DR
 DR P-PSDB; AAR12515, AAR13550.
 XX
 XX Oligo:nucleotide(s) specific for *Helicobacter pylori* - used as probes and
 PT primers to detect *H. pylori* infection, in diagnosis of gastritis, and
 PT duodenal and peptic ulceration.
 PT
 XX Disclosure; Page 16; 28pp; English.
 PS
 XX The DNA is a 2.7 kb TaqI fragment encoding the A and B subunits of *H.*
 CC *pylori* (previously *C. pylori*) urease, i.e. the 66 and 31 kD antigens.
 CC From the sequence, probes and primers can be designed for the
 CC amplification (by PCR) of the gene, to produce a prods. common to all *H.*
 CC *pylori* strains so far tested and which do not occur in other ureases and
 CC can therefore be used as a specific indication of the presence of *H.*
 CC *pylori*. *H. pylori* (previously *C. pylori*) is strongly implicated in the
 CC pathogenesis of gastritis and duodenal and peptic ulceration in man. The
 CC primers/probes can be used for the detection of *H. pylori* DNA in gastric
 CC mucosa, saliva or faecal samples to provide an early diagnosis of
 CC infection. See also AAQ12434-Q12486. (Updated on 27-AUG-2003 to correct
 CC OS field.)
 CC
 XX Sequence 2767 BP; 867 A; 550 C; 635 G; 715 T; 0 U; 0 Other;
 SQ
 Query Match 34.98; Score 1004.8; DB 2; Length 2767;
 Best Local Similarity 66.28; Pred. No. 3.8e-200;
 Matches 1627; Conservative 0; Mismatches 792; Indels 39; Gaps 10;
 QY 182 AAAATTTAACACAGAGTAAATAGGTGAAAACCTACACCCCAAGAGCAAGAAAAGTTCTTG 241
 DB 40 ACACTTTAAAGATAGGAGATGAGATGAACTACCCCAAGAGTTAGACAAAGTTGATG 99
 QY 242 TTATATTTATCGGCGGAAGTGGCTAGAAAGCGCAAGCAAGGGCTTTAAAGCTCAACAA 301
 DB 100 CTCCTACTGCTGAGAAATAGCTTAAAGAGCAAGAAAAGGCAATTAAGCTTAACTAT 159
 QY 302 CCCGAGCCATTGCTTACATTAGTGGCCATATTATGCAAGAGCGCCGCTGGAAGAAA 361
 DB 160 GTGGAAGCGGTACGTTTGTATTAGTGCCCATATTATGGAAGAGCGAGACGCGTGAAGA 219

QY 362 ACCGTTGCCAGCTTATGGAAGAGTGCACTCTTTTGAAGAAAAGATGAAGTAATGCC 421
 DB 220 ACTGCGGCTGAATTGATGCAAGAGGGCGCACTCTTTTAAACCGGATGATGTAGTGAT 279
 QY 422 GGGGTGGGTAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCTCTGATGTGAGAAA 481
 DB 280 GCGGTGGCAAGCATGATCCATGAAGTGGGTATTGAAGCGATGTTTCTCTGATGGACCAA 339
 QY 482 CTTGTAACTGTGAATTGGCCCATCGAACGAGATGAGCACTTCAAGCGGGCGAAGTGA 541
 DB 340 CTCGTAAACCGTGCAATACCCCTATTGAGGCAATGGTAAATATTAGTTCTCTGGTGA 396
 QY 542 TTTGGTTCGGATAAAGACATCGAGCTCAATGCAAGGCAAGAAAGTAACCGAACTTGA 601
 DB 397 TTTCTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAAAGCCGTTAGCGTGA 456
 QY 602 ACTAATGAAGGCGCTAAATCTTTGATGTGGGTAGCCATTTCCACATCTTTTGAAGCT 661
 DB 457 CCCCCTGTTGGGACAGACCGGTTCAATCGCTCACACTTCCATTTCTTTTGAAGTGA 516
 QY 662 AAGGCATTAATAATTCGATCGTGAAGAAAGCCTATGCAAAAGCGCTAGATATTCCTCT 721
 DB 517 AGATGCTTTAGACTTTGACAGAGAAAAAACTTTCCGTAAACCGCTTAGACATTTGCG 576
 QY 722 AACAGCTACGATTTGGGCGAGCAAAACCCGCAAGTGCAGTTGATTCTCTTGGTGGC 781
 DB 577 ACAGCGGTAAAGTTTGAAGCTTCGCAAGAAAAATCCGTAGAGTTGATTGACATTTG 636
 QY 782 AGTAAAAAAGTGAATTGGCATGAACGGGCTTGTGAATAACATCGCGAT---GAA 837
 DB 637 AACAGAGAACTCTTTGGATTAAACGCTTGGTTGATAGCAAGCCGATACGNAAGCA 696
 QY 838 -----TAAACA-----TAAAGCGC-----TTGACAAAGCGC-----AA 872
 DB 697 AAAATTTGCTTTACACAGAGCTAAAGAGCGTGGTTCATGCGCGCTAAAGAGCATGA 756
 QY 873 GATTT---ATCAAGTAAGAGACTCCCATGAAA---ATGAAAAACAAGAAATATGTA 926
 DB 757 TATGTAAAAACAATTTAAGGAGTAAAGAAATGAAAAAGATTAGCAGAAAAAGAAATAT 816
 QY 927 ACCTACGACCCCAAGAGCGATAAAGTGGCTTTAGGAGATACCGATCTTTTGGGCGA 986
 DB 817 ATGTATGCGCTTACTACAGGCGATAAAGTGAATTTGGGCGATACAGACTTGTATCG 876
 QY 987 GTAGAAATGACTATACCACTATGCGCAAGAACTTAAATTTGGCGCGGTAAAACTATC 1046
 DB 877 GTAGAAATGACTATACCACTTATGCTGAAGAGCTTAAATTTGGCGCGGTAAAAAC 936
 QY 1047 CGTGAAGGTATGGGTGAGAGCAATAGCCTGATGAAAAACCCCTAGATTTAGTCACT 1106
 DB 937 AGAGAAGGCATGAGGCAATCTAACAAACCTAGCAAAAGAAAGAACTGATCTAATCA 996
 QY 1107 AACCGATGATTATCGACTACACCGGATTTACAAAGCGGATTTGGGATTTAAAAAC 1166
 DB 997 AACGCTTTAATCGTGGATTTACCGCGTATTATTAAGCGGATTTGGTATTAAAGAT 1056
 QY 1167 AAAATCCATGCTTTGGCAAGCGAGAAACAGGACATGCAAGATGGCGTAAAGCCCTCAT 1226
 DB 1057 AAAATCGCTGCTGATTTGTTAAAGCGGTAAACAAAGACACGCAAGATGGCGTTAA 1116
 QY 1227 ATGTCGTGGGTGTGGGACAGAGCACTAGCAGGGGAGGTATGATTATTACCGCTGG 1286
 DB 1117 CTTAGCGTGGGTCTCTGCTACTGAGCCCTTAGCCGTGAAGGTTTGTATTAACTGCT 1176
 QY 1287 GGAATCGATTACACACCACTTCTTTCTCCAACAAATTTCCCTACCGCTTAGCCAAT 1346
 DB 1177 GGTATTGACACACATCTTCTATCTCCCGCAACAAATCTCTACAGCTTTTGAAGC 1236
 QY 1347 GGCCTTACAAACCATGTTTGGAGGGGACAGCTCTCTGTAGATGGCAAGATCGACTACT 1406
 DB 1237 GGTGTAAACAACCATGATTGGTGGCGGAATGCGCCCTGCTGATGGCACTAACGCAAC 1296
 QY 1407 ATCACTCCGGGCAAAATGGAACCTTGCAACCGCATGTTTGGCGGACAGAGAGTATTCT 1466

1297 ATCACTCAGGTAGAGAAATTTAAATTCATGCTCAGAGCGGTGAAGAAATATTCATG 1356
 1467 AATGTGGGCTTTTGGGCAAGGCAATAGTCTAGCAAAAAACAACCTTTGTAGAACAAAGTA 1526
 1357 AACTTTGGTTCTTGGCTAAAGGTAAAGCTCTTAACGATGCAAGCTTAGCCGATCAAAAT 1416
 1527 GAAGCGGCGCGATTGGTTTAAATTCATGAAGACTGGGGCAACAAACCAAGTGGGATC 1586
 1417 GAAGCTGGTGGATGGCTTTAAATTCACGAAGACTGGGGCAACCACTCTCTTCGCAATC 1476
 1587 GATCACTGCTTGAGCGCTGGCAGATGATAGATGTCGAAGTTTGTATCCACACGATACA 1646
 1477 AATCATGCGTTAGATGTGGGGAACAATACGATGTGAAGTCGCTATCCACACAGACACT 1536
 1647 GTCAATGAGCGAGGTATGTAGATGACACCCCTAAATGAACGCGCGCGCATCCAT 1706
 1537 TTGATGAGCGGTTGGTGGAGAGACACTATGCGAGCTATTCGCGAGCACTATGCAC 1596
 1707 GCTTACACATTTAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGGC 1766
 1597 ACTTACCACACTGAAGCGCTGGCGGCGGACACGCTCTGATATTTAAAGTGGCGGT 1656
 1767 GAGCTCAATATTCACCTCTCTCCACACCCCACTATTCCTATACCATTAATACGGTT 1826
 1657 GAACACAACTATCTCCCGCTTCCACTAACCCCACTATCCCTTCCCGTGAATACAGAA 1716
 1827 GCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGAT 1886
 1717 GCCGAACATGACATGCTTATGCTGTGCCACCACTTGGATAAAGCATTAAGAAGAT 1776
 1887 TTACAATTTTCTCAAGCGGTATCCGCGCGGCTCTATCGCGCTGAAGATGTGCTCCAT 1946
 1777 GTCCAGTTGCTGATTCAGAGGATTCGCGCTCAACCACTTGGCGTGAAGACACTTTGCAT 1836
 1947 GATATGGGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGTGCAGGCGAA 2006
 1837 GACATGGGGATTTTCTCAATCACTAGTCTGACTCTCAAGCGATGGGCGTGTGGGTGAA 1896
 2007 GTGATTCCTCGAATTCGCGAGCTGCGATGAAGATAAAGAAATTTGGTAAGCTTCCT 2066
 1897 GTTATCACTAGAACTTGGCAACAGCTGACAAAAATAAAGAAATTTGGCGCTTGAAA 1956
 2067 GAAGATGGCAAGATAACGATAATTTCCGATTAAAGCGCTACATCTCCAAATACACTATC 2126
 1957 GAAGAAAAGCGGATACGACAACTTCAGGATCAACGCTACTTGTCTAATAACCACTT 2016
 2127 AACCCCGCTTTGACCCAGCGGTGAGGAGTATATCGGCTCTGTGGAAGGGCAAGATC 2186
 2017 AACCCAGCGATCGCTCATGGGATTAGCGAGTATGTCGGTCTGTAGAAAGTGGCAAAAGTG 2076
 2187 GCCGACTTGGTGTGGATCTCGCTCTTTTGGCGTFAAACCCAAATCGTATCAAA 2246
 2077 GCTGACTTGTATTTGGAGTCCGCGATCTTGGTGTGAACCCAAACATGATCATCAA 2136
 2247 GGCGGTATGGTGGTCTTCTGAAATGGCGATTTCAACGCGTCTGTGCCCACTCCCCAA 2306
 2137 GGCGGTTTATCGCATTTGAGTCAAAATGGGTGATGCGAACGCTTCTATCCCTACCCCAA 2196
 2307 CCGGTTTATTAACCGCAATTTTGGGCATCACGCAAGGCGAAATTTGACACAGCATC 2366
 2197 CCAGTTTATTACAGAGAAATGTCGCTCATCATGTTAAAGCTAAATACGATGCAAAACATC 2256
 2367 ACTTTGTTCCAAAGTCGCTATGAATAATGGCGTGAAGAAAGCTGGGCTTAGAGCGC 2426
 2257 ACTTTTGTCTCAAGCGGCTTATGACAAAGGCAATTAAGAAAGAAATTAGGCGCTTGAAGA 2316
 2427 CAAGTTCTACCGGTCAAAACTGCGTAAACATCACCAAGAAAGACTTCAAGTTCACACAC 2486
 2317 CAAGTGTTCGCGGTAAATAATTTGAGAAATATACATAAAGAAAGACATCAATTTCAACGC 2376
 2487 AAAACGGCAAAATCACCGTGCATCCGAAAAACCTTTCGAGGTCTTTGTAGATGGCAAACTC 2546

Db 2377 ACTACCGCTCATTGAAGTCAATTTGAACTTACATGTGTCTGTGATGCAAA-GA 2435
 Qy 2547 TGCACCTTAACCCACCTCGCAAGTGCCTTAGCCAGCGCTACACTTTTCTTAGG 2604
 Db 2436 AGTAACTCTAAACCCAGCCCAATAAAGTGAAGCTTGGCGCAACTCTTTAGCATTTTCTTAGG 2493
 RESULT 14
 ID ADQ37850 standard; DNA; 1710 BP.
 AC ADQ37850;
 XX 07-OCT-2004 (first entry)
 DT H. bizzozeronii ureB DNA.
 DE Urease; urease gene cluster; urease structural gene;
 KW urease accessory gene; ureB; Helicobacter bizzozeronii infection;
 KW antibacterial; gene; ds.
 OS Helicobacter bizzozeronii.
 XX Key Location/Qualifiers
 FH CDS 1..1710
 FT /*tag= a
 FT /product= "H. bizzozeronii ureB"
 PN US2004142343-A1.
 XX 22-JUL-2004.
 PD 12-AUG-2003; 2003US-00639273.
 PF 16-AUG-2002; 2002US-0404337P.
 PR (CHAN/) CHANG Y.
 PA (SIMP/) SIMPSON K W.
 PA (ZHUJ/) ZHU J.
 XX Chang Y, Simpson KW, Zhu J;
 PI WPI; 2004-533502/51.
 DR P-PSDB; ADQ37851.
 XX Novel isolated nucleic acid molecule having urease gene cluster, and
 PT conferring on Helicobacter bizzozeronii ability to produce urease, useful
 PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
 PS Claim 7; SEQ ID NO 4; 40pp; English.
 CC The invention relates to an isolated nucleic acid molecule conferring on
 CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
 CC acid molecule is a urease gene cluster comprising at least one urease
 CC structural gene and at least one urease accessory gene. The nucleic acid
 CC molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
 CC invention also relates to an isolated protein encoded by the nucleic
 CC acid, a vaccine for preventing onset of disease in mammals infected by H.
 CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
 CC antibody or its binding portion raised against the nucleic acid. The
 CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
 CC against onset of disease caused by infection of H. bizzozeronii, which
 CC involves administering the sequences. The sequences are useful for
 CC detecting H. bizzozeronii in a sample of tissue or body fluids which
 CC involves providing a nucleic acid as an antigen, providing an antibody,
 CC or providing a nucleotide sequence as a probe in a nucleic acid
 CC hybridisation assay, contacting the sample with the antigen or the probe,
 CC and detecting any reaction which indicates that H. bizzozeronii is
 CC present in the sample. This sequence represents H. bizzozeronii ureB DNA.
 SQ Sequence 1710 BP; 492 A; 436 C; 386 G; 396 T; 0 U; 0 Other;
 Query Match 32.7%; Score 940.8; DB 12; Length 1710;

Best Local Similarity 72.0%; Pred. No. 8.3e-187; Matches 1227; Conservative 0; Mismatches 477; Indels 0; Gaps 0;	
Qy	900 AAAATGAAAAAACAAGAAATATGTAAATACCTACGAGCCCAACAAAGGCGATAAAGTGGCG 959
Ds	7 AAAATCTCTCGAAAGAAATATGTTTCTATGTATGGACCCACTACCGGCGATAAAGTGAGA 66
Qy	960 TTACGAGATCCGATCTTTGGGCGAGAGTAGACATGACTATACACCTATGGCGCAAGAA 1019
Ds	67 TTGGGCGATACCGACCTGATCTTAAAGATGCAATGACTGACCACTTATGGCGAAGAA 126
Qy	1020 CTTAAATTTGGCGGGTAAACCTATCGTGAGGGTATGGGTACAGCAATAGCCCTGAT 1079
Ds	127 ATTAAGTTTGGTGGCGGTAAACCTATCGCGATGGATGGACCAACCAAGCCCGACG 186
Qy	1080 GAAACACCTTAGATTTAGTCACTAAACGCGATGATTTGACTACACCGGGATTTAC 1139
Ds	187 AGCCACGAACTCGATCTTGCTCACTAAACGCGCTGATCGTGGATTAACCGGCTTTAT 246
Qy	1140 AAAGCCGACATTTGGATTAACCGGCAAAATCCATGGCATTTGGCAAGGAGGAAACAG 1199
Ds	247 AAAGCCGATATTTGGATTAACCGGCAAAATCCATGGCATTTGGCAAGGAGGCAATAAA 306
Qy	1200 GACATGCAAGATGGCGTAAGCCCTCATATGCTGTGGGTGGGCGACAGAGCACTAGCA 1259
Ds	307 GACATGCAAGATGGCGTTTGAACAATCTTTGGGTGGGCGCTGCTACTGAGGCTTTGGCC 366
Qy	1260 GGGGAAGGTATGATTTATCCGCTGGGGGAATGATTTACACACCCACTTCTCTTTCTCCA 1319
Ds	367 GCTGAAGGGCTGATTTTACAGCTGGTGGGATTCACACCCACATCCACTTTATTTCTCCC 426
Qy	1320 CAACAATCCCTACCGCTTAGCCAAATGGCGTTTAAACATGTTTGAAGGCGGCAAGGT 1379
Ds	427 CAACAATCCCGACAGCATTTGGCAGCGGGATCAACACCATGATTTGGTGGGGCAACAGGT 486
Qy	1380 CTTGTAGATGGCGATGCGACTACTATCACTCGGGCAAAATGGAATTTGCAACGATG 1439
Ds	487 CCAGCTGATGGGACTAAGCGGACTACCATCACTCGGGCGCGCTGGAACTTTAAACACCATG 546
Qy	1440 TTGCGCGACGAGAAGATTTCTTAATGTTGGCTTTTGGGCAAGGCAATAGCTCT 1499
Ds	547 CTGCGTCTCTGAGNATATGCAATGAACTTTGGGCTATTTGGTAAAGGAAATGTGTCT 606
Qy	1500 AGCAAAAAACAATTTGAGAACAAAGTAGAAGCGGGCGCGATTTGTTTTAAATTCATGAA 1559
Ds	607 TATGAACCTCTCCCTGGTTCGATCAACTCGAAGCTGGAGCCATTTGGCTTTAAATCCACGAA 666
Qy	1560 GACTGGGCGCAACACCAAGTCGATGCTGCTTTGAGCGTGGCGAGATGAATACGAT 1619
Ds	667 GACTGGGCGTAGCACCTCGAGCGCATCTACCATTTGCTTGAATGTGGCTGACAAATACGAT 726
Qy	1620 GTGCAAGTTTGTATCCACCGGATACAGTCAATGAGGCGAGTTATGTAGATGACACCTTA 1679
Ds	727 GTGCAAGTTGCTATCCACCGGATACCTTGAATGAGCGGGCTGTGGGAGACACTTTG 786
Qy	1680 AATGCAATGAAAGCGGCGCGCATCTCATGCTTACACATTTGAGGAGCGGCTGGAGGACAC 1739
Ds	787 CAAGCCATTTGCTGGCGCACTATCCACACTTTCCACACTGAGGTGCTGTGGCGGCGAC 846
Qy	1740 TCACCTGATGTTATACCATGCGAGGAGCTCAATTTTACCTCTCCACACCCCGCC 1799
Ds	847 GCTCCGAGTGTCAATTAAGATGTCTGGCAATTTAAACATCTCCAGCTTTACCAACCCCG 906
Qy	1800 ACTATTCCCTATACCATTAATAGGTTGCAAGACACTTAGACATGCTCATGACATGCCAC 1859
Ds	907 ACCATTCTCTTCCCGTGAATCAGAAAGCCGAAACACATGACATGTTGATGGTGTGCCAC 966
Qy	1860 CACTAGACAAACGCTCCGCGAGGATTTACAAATTTTCTCAAGCCGTATCCGCGCCCGGC 1919
Ds	967 CACTTGGATTAACAACTCAAGAGATGTCAGTTGCTGATTTCTAGGATTTGCGCCCGCAA 1026
Qy	1920 TCTATCCGCGCTGAAGATGCTTCCATGATATGGGTGTGATCGCGATGCAAGCTCGGAT 1979

Ds	1027 ACCATCGCGCTGAGGACAAACTCCAGGATATGGGATTTTCTCTATCACCAGCTCTGAC 1086
Qy	1980 TGGCAAGCAATGGGCGTGCAGGCGAAGTGTATCTCGAAGTCTGGCAGACTCGGAGTAAG 2039
Ds	1087 TCCCAAGCGATGGGCGTGTAGCGAGGTATCACCGGCACTTTGGCAAAACAGCGGACAAA 1146
Qy	2040 AATAAAAAAGAAATTTGGTAAAGCTTCTGAAGATGGCAAGATACCAAGATAATATTCGCGATT 2099
Ds	1147 AACAAAAAGAAATTTGGTTCGCTGAGGAAAAAGGCGATATGACAACTTCCGATC 1206
Qy	2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTAT 2159
Ds	1207 AAGCGCTACATTTCCAAATACACTATCAACCCCGCTTTGACCAACGCGCACTTTCTGAATAT 1266
Qy	2160 ATCGGCTCTGTGGAAGAGGCGAAGATCGCGACTTGGTGGTGTGGAATCTCGCTTTTTT 2219
Ds	1267 GTCGGCTCTGTAGAAAGTGGGCAAAATTCGCGGATTTGGTGTGGAGTCTCGGTTCTTT 1326
Qy	2220 GGCGTAAACCCCAAAATCGTGATCAAGGCGGTATGGTGTCTTCTGAAATGGGCGAT 2279
Ds	1327 GGCATTAACCCCAACATGATCAAGGCGGATTCATCGCACTTTCTCAATGGGCGAT 1386
Qy	2280 TCTAACGGCTGTGTCCTCACTCCCAACCGGTTTATTACCGGAAATGTTTGGCGATCAC 2339
Ds	1387 GCGAATGCTCTATCCCACTCCCAACCGGTTATTACCGGAAATGTTTGGCCCAT 1446
Qy	2340 GGCAGGCGAAATTTGACACCGATCACTTTTGTTCAAAGTCGCTATGAAATGGC 2399
Ds	1447 GGTAAAGCCAAATTTGACACCAATATCACTTTTGTATCCCAAGTGGCTTATGACACGCG 1506
Qy	2400 GTGAAGAAAGCTGGGCTTAGAGCGGCAAGTTCTACCGTCAAAAATGCGCGTAACATC 2459
Ds	1507 ATTAAGAAAGTGGGCTTGCAAGAGTGGTTCGCAAGTGTTCGCAAGTGTTCGCAACATC 1566
Qy	2460 ACCAAGAAAGCTTCAAGTTCAACCAACCAAAACGCAAAATCACCGTCCGATCGAAGAAC 2519
Ds	1567 ACCAAAAAGACCTCAAAATCAACGATGTACCGCACATCGAAGTCAATCTCGAAGAAC 1626
Qy	2520 TTGCGAGTCTTTGTAGATGGCAAACTCTGCACTCTAAACCCACCTTCGCAAGTGCCTCTA 2579
Ds	1627 TACAAGTTAAAGTGGATGGCAAGAGGTACTTCCAAAGCAGCGGATAAATCAGCCTA 1686
Qy	2580 GCCAGCGCTACACTTTCTCTAG 2603
Ds	1687 GCACAACTTACAACTTCTCTAG 1710
RESULT 15	
ABA00816	
ID	ABA00816 standard; cDNA; 1719 BP.
XX	AC ABA00816;
XX	AC ABA00816;
DT	01-APR-2003 (first entry)
XX	H. felis urease B coding sequence.
DE	Gene; rat; IgG2a; light chain; E. coli; galactosidase; immunogen;
KW	immune response; circulatory vessel; Gut Associated Lymphoid Tissue;
KW	GALT; immune response; antigen; Salmonella; Cholera; immunity;
KW	Helicobacter pylori; HIV; Candida; P. gingivalis; gut; parasite; toxin;
KW	hormone; hormone receptor; cancer; ss.
XX	Helicobacter felis.
OS	WO200296949-A1.
FX	05-DEC-2002.
PD	24-MAY-2002; 2002WO-AU000661.
PF	25-MAY-2001; 2001AU-00005241.
XX	
XX	

PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
 XX Mckenzie B, Boyle J, Lew A;
 XX WPI; 2003-156789/15.
 XX Raising an immune response in an animal by administering composition
 PT comprising carrier and antigen bound to targeting moiety which binds to
 PT receptor present in circulatory vessels in Gut Associated Lymphoid
 PT tissue.
 XX Disclosure; Page 27-28; 45pp; English.
 XX This sequence encodes H. felis urease B. This sequence was used in the
 CC method of the invention for raising an immune response in an animal. The
 CC method comprises administering to the animal a composition comprising a
 CC carrier and an antigen bound to a targeting moiety which binds to at
 CC least one receptor present in circulatory vessels in Gut Associated
 CC Lymphoid Tissue (GALT). The method is useful for raising an immune
 CC response in an animal against antigens from Salmonella, Cholera,
 CC Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut
 CC associated toxins, gut hormones, gut hormone receptors or gut associated
 CC cancers. The method is useful for raising both mucosal and systemic
 CC immunity against any antigen used in the composition
 XX
 SQ Sequence 1719 BP; 490 A; 411 C; 421 G; 397 T; 0 U; 0 Other;
 Query Match 31.3%; Score 902.4; DB 8; Length 1719;
 Best Local Similarity 70.6%; Pred No. 9.1e-179;
 Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;
 900 AAAATGAAAAAACAAGATATGTAATACCTACCGACCCACCAAGGCGATAAAGTGGCG 959
 15 AAGATTTACGAAAGAAATATGTTCTATGATGTGTCCTACCTACCGGGATCGTGTAGA 74
 960 TTAGAGATACCGATCTTTGGGCGAAGTAGAACAATGATATACCACTATGGCGAAGAA 1019
 75 CTCGGCGACACTGATTTGATCTTAGAAGTGGAGCATGATGACCACTTATGTTGTAAGAG 134
 1020 CTTAAATTTGGCGGGTAAACTATCCGTGAGGTATGGGTACAGCAATAGCCCTGAT 1079
 135 ATCAAAATTTGGGGCGGTAACATATCCGTGATGGATGAGTCAAAACCAATAGCCCTAGC 194
 1080 GAAAAACCCCTAGATTTAGTCATCACTAACCGCATGATTTATCGACTACACCGGGATTTAC 1139
 195 TCTTATGAAATTTAGTTGCTGCTCACTAACGCCCTCATTTGGACTATACGGGCATTTAC 254
 1140 AAAGCGCATTTGGATTAATAACCGGCAAAATCCATGGCATTTGGCAAGCAGGAAACAAG 1199
 255 AAAGCGGACATTTGGATTAAGACCGGCAAGATTGCGAGGCATTTGGCAAGGCGCAATAAG 314
 1200 GACATGCAAGATGGGTAAAGCCCTCATATGTCGTGGGTGGGCGACAGCAAGCACTAGCA 1259
 315 GACATGCAAGATGGGTAGATAAATATCTTTGCGAGTCTGCTACAGAGGCTTTGGCA 374
 1260 GGGGAAGTATGATTAATACCGCTGGGGAATCGAATTCACACACCCACTTCTCTTCTCCA 1319
 375 GCTGAGGCTTTGATTTGAACCGCTGGTGGCATCGATACGATATTCATCTTCTCTCC 434
 1320 CAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTACAACCAATGTTTGGAGCGGCGA 1379
 435 CAACAATTTCCCTACTGCTTTTGGCAGCGGGTTACAACCAATGTTTGGAGGAGGCA 494
 1380 CTTGATGATGCGAAGATGCGACTACTATCACTCCGGGCAAAATGGAATTTGCGACCGATG 1439
 495 CTTGGATGCGAAGATGCGAAGATGCGAAGATGCGAAGATGCGAAGATGCGAAGATG 554
 1440 TTGCGCGCAGCAGAGATGATTTCTATGATGTTGGCTTTTGGGCAAAAGCAATAGCTCT 1499
 555 TTGCGCGCAGCAGAGATGATTTCTATGATGTTGGCTTTTGGGCAAAAGCAATAGCTCT 614
 1500 AGCAAAAAACAACCTTTGTAGAACCAAGTAGAAGCGGGCGGCGATTTGTTTAAATTTGCA 1559

615 TAGGAACCCCTCTTTAGCGCATCAGATTGAAGCAGGGCGGATTTGTTTAAATCCAGAA 674
 1560 GACTGGGGCACAACCAACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
 675 GACTGGGGAAGCAGACACCTGCGAGCTATTACCACTGCTCAATGTCGCCGATGAATACGAT 734
 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGTTATGTAGATGACACCCCTA 1679
 735 GTGCAAGTGGCTATCCACACCGATACCTTAAACGAGCGGGCTGTGTAGAGACACCCCTA 794
 1680 AATGCAATGAACGGGCGGCCCATCCATGCTTACCAATGAGGAGCGGGTGGAGACAC 1739
 795 GAGGCGATTTGCGCGCGGCACCATTCATCTTCCACATGAGGGGCTGGGGTGGAGACAC 854
 1740 TCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCCCTCTCCACACCCCCC 1799
 855 GCTCCAGATGTTATCAAAATGGCAGGGAATTTAACTTTACCCGCTCTCTACTAACCCG 914
 1800 ACTATTCCCTATACCAATTAATAGGTTGAGAAACACTTAGACATGCTCATGACATGCGAC 1859
 915 ACCATTCTCTTTCACCAAAACACTGAAGCGGAGCAGCATGGACATGTTAATGGTGTGCGAC 974
 1860 CACCTAGACAAACGCGATCCGCGAGGATTTACAATTTTCTCAAAGCGTATCCCGCCCGGC 1919
 975 CACTTGGATAAAGTATCAAGGAAGATGTGCGATTTGCCGATTCGAGGATTCGCCCCCAA 1034
 1920 TCTATCGCGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGGAT 1979
 1035 ACTATCGCGCTGAAGACCAACTCCATGATGATGGGATCTTTTCTATCACCAGCTCCGAC 1094
 1980 TCGAAGCAATGGGCGTGCAGCGAAGTATTCCTCGAACTTTGGGAGACTGGGATGAAG 2039
 1095 TCTCAGGCTATGGGACGCGTATGGCGAGGTGATCACACGCACTTTGGCAGACAGACAAA 1154
 2040 AATAAAGAAATTTGGTAAGCTTCTCTGAAGATGGCAAGATTAACCAATATTTCCCGATT 2099
 1155 AAAAAAGAGTTTGGGCGCTTGAAGAGAAAAAGGCGATTAACGACAACTTTCCGATC 1214
 2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGCGTGAGCGATAT 2159
 1215 AAGCGTACATCTTAATACACCATCAACCCCGCGATCGCGCATGGGATTTCTGACTAT 1274
 2160 ATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTGGTGTGGAATCTCGCTTTTCTT 2219
 1275 GTGGGCTCTGTGGAAGTGGCAAAATACGCGACCTCGTGTCTTTGGAGTCCGGCTTTCTT 1334
 2220 GGGTAAAAACCAATCGTATCAAGGCGGTATGGTGTCTTCTCTGAAATGGGCGAT 2279
 1335 GGCATTAAGCCCAATATGATTTAAGGGCGGATTTATTGCGCTCTCTCAATGGGCGAT 1394
 2280 TCTAAGCGCTCTGTGCCCACTCCCAACCGGTTTATTACCGCGAAATGTTTGGGCGATCAC 2339
 1395 GCAATGCGCTTATTTCCACCCCTCAGCCGCTCTATTACCGTGAATGTTTGGACACAT 1454
 2340 GGCAGGCGGAAATTTGACACCAAGCATCACTTTTGTGTTTCCAAAGTCGCTATGAAATGGC 2399
 1455 GGGAAAAACAAATTCGACACCAATATCACTTTCTGTGTCCTCAAGCGGCTTACAAGCGAGG 1514
 2400 GTGAAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAAATGCGCGTAAATC 2459
 1515 ATCAAGAGAGAACTAGGGCTAGATTCGCGTGTGTTGCGAGTGAATAAACTGTGCGCAATATC 1574
 2460 ACCAAGAAAGACTTCAAGTTTCAACGACAAACCGGCAAAAATCAACCGTCAATCCGAAAC 2519
 1575 ACTAAAGAGACCTCAATTTCAACGATGTGACCGCATATGATGATCAACCTTGAAC 1634
 2520 TTGAGGCTCTTTGTAGATGGCAAACTCTGACCTCTAAACCCCACTCGCAAGTGCCTCTA 2579
 1635 TATAAGTGAAGTGGTGAAGAGAGTAACTCTTAAAGCAGCAGATGAATGAGCCTA 1694
 2580 GCCAGCGCTACACTTTCTCTAG 2603
 1695 GCGCAACTTTATAATTTGTTCTAG 1718

Search completed: August 10, 2006, 01:43:00
Job time : 1755 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:37:34 ; Search time 528 Seconds
(without alignments)
10216.679 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2880.6

Sequence: 1 rgragattttccaractt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY NUC

Gapop 3.0 , Gapext 3.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /EMC_Celerra_SIDS3/prodata/2/ina/1_COMB.seq:*
- 2: /EMC_Celerra_SIDS3/prodata/2/ina/5_COMB.seq:*
- 3: /EMC_Celerra_SIDS3/prodata/2/ina/6A_COMB.seq:*
- 4: /EMC_Celerra_SIDS3/prodata/2/ina/6B_COMB.seq:*
- 5: /EMC_Celerra_SIDS3/prodata/2/ina/7_COMB.seq:*
- 6: /EMC_Celerra_SIDS3/prodata/2/ina/8_COMB.seq:*
- 7: /EMC_Celerra_SIDS3/prodata/2/ina/PCUS_COMB.seq:*
- 8: /EMC_Celerra_SIDS3/prodata/2/ina/PP_COMB.seq:*
- 9: /EMC_Celerra_SIDS3/prodata/2/ina/RE_COMB.seq:*
- 10: /EMC_Celerra_SIDS3/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1021.4	35.5	2735	2	US-08-920-095-1
2	1021.4	35.5	2735	7	PCT-US96-05800-1
3	1008.8	35.0	2619	2	US-08-467-822-19
4	1008.8	35.0	2619	3	US-08-432-697-19
5	1008.8	35.0	2619	3	US-08-466-248-19
6	1005.8	34.9	4824	3	US-09-431-705-1
7	1005.8	34.9	4824	3	US-09-431-705-19
8	667.4	23.2	6131	2	US-07-732-242C-8
9	665.4	23.1	1830121	3	US-09-557-884-1
10	665.4	23.1	1830121	3	US-09-643-990A-1
11	665.4	23.1	1830121	3	US-10-158-865-1
12	648.2	22.5	1710	3	US-09-543-681A-1857
13	617.2	21.4	2400	2	US-08-967-513-1
14	617.2	21.4	2400	2	US-08-687-645B-1
15	612.8	21.3	1878	3	US-09-489-039A-2045
16	610.2	21.2	1716	3	US-08-487-429A-3
17	610.2	21.2	1716	7	PCT-US96-05320A-541
18	593.6	20.6	5966	3	US-08-956-171B-22
19	593.6	20.6	5966	3	US-08-781-986A-22
20	573.4	19.9	1704	3	US-09-328-352-1786
21	569.8	19.8	8729	3	US-09-453-702B-258
22	569.8	19.8	8729	3	US-10-114-170-258
23	569.8	19.8	87563	3	US-09-453-702B-57

24	569.8	19.8	87563	3	US-10-114-170-57	Sequence 57, Appl
25	567.6	19.7	1701	3	US-09-252-991A-10316	Sequence 10316, A
26	563.8	19.6	3164	3	US-09-710-279-3828	Sequence 3828, Ap
27	563.8	19.6	3234	3	US-09-710-279-3853	Sequence 3853, Ap
28	562.8	19.5	1716	3	US-09-710-279-117	Sequence 117, App
29	562.8	19.5	1722	3	US-09-134-001C-2189	Sequence 2189, Ap
30	552	19.2	1653	3	US-09-252-991A-10546	Sequence 10546, A
31	542	18.8	1833	3	US-09-602-777A-13	Sequence 13, Appl
32	518.6	18.0	1625	3	US-09-602-777A-15	Sequence 15, Appl
33	473.8	16.4	1797	3	US-09-601-198-89	Sequence 89, Appl
34	470.6	16.3	4403765	3	US-09-103-840A-2	Sequence 1, Appl1
35	470.6	16.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
36	374.2	13.0	3475	3	US-09-710-279-4032	Sequence 4032, Ap
37	356.6	12.4	951	3	US-09-252-991A-10513	Sequence 10513, A
38	211.2	7.3	684	3	US-10-012-819-153	Sequence 153, App
39	208.2	7.2	636	3	US-09-252-991A-10653	Sequence 10653, A
40	173.2	6.0	363	3	US-10-012-819-159	Sequence 159, App
41	106.8	3.7	462	3	US-09-543-681A-1913	Sequence 1913, Ap
42	106.6	3.7	613	3	US-09-533-559-7520	Sequence 7520, Ap
43	100.2	3.5	312	3	US-09-328-352-1761	Sequence 1761, Ap
44	99.6	3.5	360	3	US-09-328-352-1789	Sequence 1789, Ap
45	98.8	3.4	399	3	US-09-134-001C-2212	Sequence 2212, Ap

ALIGNMENTS

RESULT 1
US-08-920-095-1
; Sequence 1, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-920-095-1

Query Match 35.5%; Score 1021.4; DB 2; Length 2735;
Best Local Similarity 66.5%; Pred. No. 6.6e-244;
Matches 1634; Conservative 0; Mismatches 786; Indels 38; Gaps 9;

Qy	182	AAAATTAAACACAAGGAGTAAATAGGTGAACCTCACACCCCAAGACGACGAAAAAGTTCTTGT	241
Db	78	ACACTTTAAGAAATAGGAGAAATGAGATGAACACTCACCCCAAAAGAGTTAGATAAAGTTGATG	137
Qy	242	TTATATTATTCGGCGGCAAGTGGCTAGAAAACGCGAAAGCAGACGGCTTTAAAGCTCAACCCA	301
Db	138	CTCCACTACCTCGAGAAATTTGGCTTAAAAAACGAAAGAAAAGGCATTTAAAGCTTAACTAT	197
Qy	302	CCGGAAGCCATTGCTTACATTAGTGGCCCATATTATGGACGAAGCGCGCCGTGGAAAAAAA	361
Db	198	GTAGAAGCAGTAGCTTTGATTAGTGGCCCATATTATGGAAGACGAGAGCTGGTAAAAAG	257
Qy	362	ACCGTTCCCGAGCTTATGGAAGAGTGCATGCACTTTTGTAAAAAAGATGAAGTAATGCCCC	421
Db	258	ACTGCGGCTCGAATTGATGCAAGAAGGCGCACTCTTTTAAACACGATGATGTGATGGAT	317
Qy	422	GGGTGGGTAAATATGGTTCCCGATCTAGTGTAGAGCCACCTTCTCTGATGTGTCGAAA	481
Db	318	GGCGTGCAAGCATGATCCATGAAGTGGGTATTGAAGCGATGTTTTCTGATGGGACTAAA	377
Qy	482	CTTTGTAACTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAAGCGGCGCAAGTGAAA	541
Db	378	CTCGTAAACCGTGCATACCCCTATTTGAGGCGCAATGGTAAATTTAGTTCCTGCTGAGTTG	434
Qy	542	TTTGGTTGCGATTAAGACATCGAGCTCAATGCGAGGCGAAAGAAAGTAACCGCAACTTGAGTT	601
Db	435	TTCTTAAAAAATGAAGACATCACTATCAACGAAGGCCAAAAAAGCCGTTAGCGTGAAAGTT	494
Qy	602	ACTAATCAAGCGGCTAAATCTCTTGATGTGGGTAGCCATTTCACATTTCTTTTGAAGCTAAC	661
Db	495	AAAAATGTTGGCGACAGACCGGTTCAATCGGCTCACATTTCTTCTTTGAAGTGAAT	554
Qy	662	AAGSCACTAAAAATTCGATCGTGAAAAAGCCTATGGCAACCGCTAGATATTCCTCTGGC	721
Db	555	AGATGCTAGACTTTTGACAGAGAAAAAATCTTCGGTAAACGCTTAGACATTCGCGAGCGG	614
Qy	722	AACAGCTACGCAATGGGGCAGGACAAAACCGGCAAAGTGCAAGTTGATTCCTCTTGGTGGC	781
Db	615	ACAGCGGTAGATTTTGAGCTCTGGCGAAGAAAAATCCCGTAGAAATTTGATTCGATTTGGCGGT	674
Qy	782	AGTAAAAAAGTGATTGGCATGAACGGGCTTGTGAATAACATCGCGG---ATGAACGCGCA-	837
Db	675	AACAGAAGATCTTTGGATTTAACGCATTTGGTTGATAGACAGCAGACACGAAAGCAAA	734
Qy	838	-----TAAACA-----TAAAGCGC-----TTGACAAGCGG-----AA---ATCTCAAG	872
Db	735	AAATTTGCTTTACACAGAGCTAAAGACGCTGGTTTTTCATGCGCTAAAGACGATGACAAC	794
Qy	873	GATTT---ATCAGTAGGAGACTCCCATGAAA---ATGAAAAAACAAGAAATATGTAAAT	926
Db	795	TATGTAAAAACAATTAAGAGTAGTAAGAAATGAAAAAGATTAGCGAAAGAAATATGTTTCT	854
Qy	927	ACCTACGGAACCCAAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTTGGGCGAGAA	986
Db	855	ATGTATGGTCTACTACAGCGGATTAAGTGAATTTGGCGATACAGACTTGAATCGCTGAA	914
Qy	987	GTAGAACATGACTATACCACTATGGCGAAGAACTTAAATTTTGGCGGGGTAAAACTATC	1046
Db	915	GTAGAACATGACTACACCAATTTATGGCGAAGAGCTTAAATTCGGTGGCGGTAAAAACCTA	974
Qy	1047	CGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTTAGTCACTCACT	1106
Db	975	AGAAAGGCATAGGCCAATCTAACACCCCTAGCAAGAGAGAGTTTGAATTTAAATATCACT	1034
Qy	1107	AACCGCATGATTATCGACTACCGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGC	1166
Db	1035	AACGCTTTAATCGTGGATTACACGGGTATTTTATAAGCGGATATTGTTATTAAGATGGC	1094
Qy	1167	AAANTCATGGCATTTGGCAAGGCAAGGAAAAAAGGACATGCAAGATGCGGTAAAGCCCTCAT	1226
Db	1095	AAAATCGCTGGCATTTGGTTAAAGGCGGTAAACAAAGACATGCAAGATGCGTTAAAAACAAT	1154

Qy	1227	ATGGTCTGTGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGTATGATTTATTAACGCTGGG	1286
Db	1155	CTTAGCGTAGGTCTCTGTACTGAAGCGTTTAGCCGGTGAAGGTTTGATCGTAAACGGCTGGT	1214
Qy	1287	GGAAATCGATTTCACACACCCACTTCTTTCTCCACAACAATCCCTACCGCTCTAGCCAAT	1346
Db	1215	GGTATGACACACATCTCCACTTCATTTTCACCCCAACAATCCCTACAGCTTTTTCAGAC	1274
Qy	1347	GGCGTTTAAACCATTTTGGAGCGGCACACAGGTCTCTGTAGATGGCAGTAATGGACTACT	1406
Db	1275	GGTGTAAACCATGATTGGTGGTGAACCGGTCTCTGCTGATGGCACTAATGGCACTACT	1334
Qy	1407	ATCACTCCGGGCAAAATGGAACTTTGCHACCGCATGTTGGCGGCACAGAGAGTATTTCTATG	1466
Db	1335	ATCACTCCAGGCAGAAATTTTAAATTTGAATGCTCTCAGAGCGGCTGAAGAATATTTCTATG	1394
Qy	1467	AAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAACAACTTGTAGAACAGTA	1526
Db	1395	AAATTTAGTTCTTTGGCTAAAGGTAAACGTTCTTAAAGATGGAGCTTAGCCGATCAAATT	1454
Qy	1527	GAAGCGGGCGGATTTGGTTTTTAAATTCATGAAGACTGGGGCAACAACCAAGTCGCAATC	1586
Db	1455	GAAGCGGTGGATTTGGCTTTTAAATTCACGAAGACTGGGGCAACCACTCTCTCTGCAATC	1514
Qy	1587	GATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATTCACACCGATACA	1646
Db	1515	AATCATCGTTAGATGTTGCGGCAAAATACGAATGTGCAAGTCGCTATPCCACACAGACACT	1574
Qy	1647	GTCAATGAGCGGTTATGTAGATGACACCTTAAATGCAATGAACGGGGCGGCATCCAT	1706
Db	1575	TTGAATGAAGCCGGTTGTGTAGAAGACATATGGCTGCTATGCTGAGCGCATATGCAAC	1634
Qy	1707	GCCTTACCACATTTAGGGAGCGGCTGGAGGACACTCACTGATGTTATCAACATGGCAGGC	1766
Db	1635	ACTTTCACACTGAAGCGCTGGCGGCGACACGCTCTGTATTAATTAAGTAGCCGCT	1694
Qy	1767	GAGCTCAATATTTACCCCTCTCCACACCCCCCACTAATTCCTTATACCATTAATACGGTT	1826
Db	1695	GAACACAACATTTCTTCGGCTTCCACTTAACCCCAACCACTCCCTTTCACCGTGAATACAGAA	1754
Qy	1827	GCAGAACACTTAGACATGTCATGACATGCCACCACCTAGACAAAAGCATCCGCGAGGAT	1886
Db	1755	GCAGAGCACATGACATGCTTATTTGGTGTGCCACCACTTGGATAAAAGCAATTAAGAAGAT	1814
Qy	1887	TTACAAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTGCTCCAT	1946
Db	1815	GTTCAGTTCCGTGATTTCAAGGATCCGCCCTCAAAACATTTGGCGCTGNAGACACTTTGCAT	1874
Qy	1947	GATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGCAA	2006
Db	1875	GACATGGGATTTTCTCAATCACAGTTCTGACTCTCAAGCGATGGSCGTGTGGGTGAA	1934
Qy	2007	GTGATTCCTCGAATTTGGCAGACTGGCGGATAGCAATAAAAAGAAATTTGTGTAGCTTCCT	2066
Db	1935	GTTATCTATGAACCTTGGCAACAGCTGACAAAACAAAGAAATTTGGCCGCTTGAAA	1994
Qy	2067	GAAGATGGCAAGATAACGATAATTTCCGCATTTAAGCGCTTACATCTCCAATACACTATC	2126
Db	1995	GAAGAAAAGCGGATAACGACAACTTCAGATCAAAAGCTATTTGTCTAAATACACCAATT	2054
Qy	2127	AACCCCGCTTTGACCCACCGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATC	2186
Db	2055	AAACCCGGATCGCTCATGGGATTTAGCGAGTATGAGTTTCAGTAGAAGTGGGCAAAAGTG	2114
Qy	2187	GCGCACTTGGTGTGGAAATCTCTGCGCTTTTTTTGGGCGTTAAACCGAAAATTCGTGATCAAA	2246
Db	2115	GCTGACTTGGTATTTGTGGAGTCCAGCAATTTCTTTGGGCTGAAACCCCAATGATCATCAAA	2174
Qy	2247	GGCGGATGTGTGTTCTCTCTGAATTTGGGCGATTTCTTAAACGCGTGTGCGCCACTCCCCAA	2306
Db	2175	GGCGGATTCATTTGGGTTAAGCCAAATTTGGGCGATGCGAACGCTTCTTATCCCTTACCCCAAA	2234
Qy	2307	CCGGTTTTTATACCGCGAAATTTGTTGGGCATCAACGGCAAGGCGAAATTTGACACCCAGCATC	2366

1287 GGAATCGATTTCACACCCCACTTCTTCTTCCCAACAAATTCCTACCGCTCTAGCCAAAT 1346
1215 GGTATTGACACACACATCCCACTTCAATTTCCACCCCAACAAATCCCTACAGCTTTTGGCAAGC 1274
1347 GCGTTACACCACTGTTTGGAGGGGACAGGTCCTGTAGATGGACGAAATGGCACTACT 1406
1275 GGTGTAAACCACTGATTTGGTGGTGAACCGGTCCTGTATGGCACTTAATGGCACTACT 1334
1407 ATCACTCCGGGCAAAATGGAACTTCACCGGATGTTGGCGCAGCAGAGAGATATTCATG 1466
1335 ATCACTCCAGCAGAAATTTAAATGGATGCTCAGAGCGGCTGAAGAATATTCATG 1394
1467 AATGTGGCTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAACTTGTAGAACAAAGTA 1526
1395 AATTTAGGTTTCTGGCTAAGGTAAGCTTCTTAACGATGCGAGCTTAGCGGATCAAAAT 1454
1527 GAAGCGGGCGGATGTTGTTTAAATTCATGAGACTGGGGCAACACCAAGTGGCATC 1586
1455 GAAGCGGGTGGATTTGGCTTTTAAATTTCAAGAGACTGGGGCACCACTCTCTTCTGCAATC 1514
1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTCATCCACACCGATACA 1646
1515 AATCATGCGTTAGATGTTGGCAAAATACGATGTGCAAGTTCGCTATCCACACAGACACT 1574
1647 GTCAATGAGCGAGGTATGATAGATGACACCCCTAAATGCAATGAACGGGCGCGCCATCCAT 1706
1575 TTGAATGAAGCGGTTGTTAGAGACACTATGCTGCTATTTGCTGGAGCACTATGCAC 1634
1707 GCCTACACATTTGAGGAGCGGTTGAGGACACTCACTGATGTTATCACCATGGCAGGC 1766
1635 ACTTTCACACTGAAGCGCTGGCGGGGACACGCTCTCTGATATTTAAATAGTAGCGGT 1694
1767 GAGCTCAATATCTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGTT 1826
1695 GAACACAACTTCTCCGCTTCCACTAACCCACCACTCCCTTCCCGTGAATACAGAA 1754
1827 GCAGAACACTTAGACATGCTCATGACATGCGCACCCACCTAGACAAACGCGATCCGGAGGAT 1886
1755 GCAGAGCACATGGACATGCTTATGTTGTCACCACTTGGATAAAGCAATTAAGAAGAT 1814
1887 TTACAAATTTCTCAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCAT 1946
1815 GTTCAGTTGCTGATTTCAAGGATCCGCTCAAAACCACTTGGCGTGAAGACACTTTGCTAT 1874
1947 GATATGGTGTATCGCGATGACAGACTCGGATTCGCAAGCAATGGCGGTGCGAGGCAA 2006
1875 GACATGGGGATTTTCTCAATCACAGTCTGACTCTCAAGCGATGGGCGGTGGGTGAA 1934
2007 GTGATTCCTCGAACTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTAAGCTTCCT 2066
1935 GTTATCACTAGACTTGGCAACAGCTGACAAACAGAAAGATTTGGCGGCTTGAA 1994
2067 GAAGATGGCAAGATAACGAATTTCCGANTTAAGCGCTACATCTCAAAATACACTATC 2126
1995 GAAGAAAAAGCGGATAACGACAACTTCAGGATCAACGCTACTTGTCTAAATACACCAT 2054
2127 AACCCCGTTTGACCCAGCGGTGAGGATATATCGGCTCTGTGGAAGAGGGCAAGATC 2186
2055 AACCCAGCGATCGCTCATGGGATTTAGCGAGTATGAGTTCCAGTAGAAGTGGGCAAGTG 2114
2187 GCCGACTTGGTGGTGGAAATCCCTGCTTTTGGCGTAAACCCAAATCGTGATCAAA 2246
2115 GCTGACTTGGTATTTGGAGTCCAGCAATTTTGGCGTGAACCCCAATGATCATCAAA 2174
2247 GCGGATATGGTGGTCTTTCTCTGAAATGGCGGATTTCAACGCGTCTGTGCCCACTCCCAA 2306
2175 GCGGATTCATTTGGTTAAGCAAAATGGCGGATGCGAAGCTTCTATCCCTACCCACAA 2234
2307 CCGGTTTATACCGCGAAATGTTTGGGATGATCAGCGAAGGGCAAAATTTGACACCGAGATC 2366
2235 CCGGTTTATTAACAGAAATGTTGCTCATCATGTTAAAGCTAAATACGATGCAAAACATC 2294

2367 ACTTTGTTTCCAAAGTCGCTATGAAATGCGGTGAAGAAAGCTGGCTTAGAGCGC 2426
2295 ACTTTTGTGTCTCAAGCGGCTTATGACAAAGGCATTAAGAAAGATTTAGGACTTGAAGA 2354
2427 CAAGTTCTACCGGTCAAAACCTGCCGTAAACATCACCAGAAAGACTTCAAGTTCAACGAC 2486
2355 CAAGTTTGGCGGTAAAAAATTCAGAAATATCACTAAAAAAGACATGCAATTCACGAC 2414
2487 AAAACGGCAAAATCACCGTCGATCCGAAAACTTTCGAGGTCTTTGTAGATGGCAAACTC 2546
2415 ACTACTGCTCACAATTGAAGTCAATCTGAAACTTACCATGTGTTCGTGGATGGCAAGAA 2474
2547 TGCACCTTAAACCCACCTCGCAAGTGCCTTAGCCGAGCGCTACACTTCTTCTTAGG 2604
2475 GTAACTTCTAAACCGCAATAAAGTGAGCTTGGCGCAACTCTTTAGCATTTTCTAGG 2532

RESULT 3
US-08-467-822-19
; Sequence 19, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36

OTHER INFORMATION: /standard_name= "Shine-Dalgarno
OTHER INFORMATION: sequence."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 756..759
OTHER INFORMATION: /standard_name= "Shine-Dalgarno
OTHER INFORMATION: sequence."
US-08-467-822-19

Query Match 35.0%; Score 1008.8; DB 2; Length 2619;
Best Local Similarity 66.2%; Pred. No. 8.9e-241;
Matches 1634; Conservative 0; Mismatches 792; Indels 42; Gaps 8;

Qy	175	TGCTTTAAATTTAAACACAGGAGTATAGTGAACCTCACCCAAAGAGCAGAAAA	234
Db	13	TACCAATAGAAATTTCA-ATAAGGAGTTTAGGATGAACCTAACGCTAAAGAACTAGACAA	71
Qy	235	GTTCCTGTATTATATGCGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCT	294
Db	72	GTTAATGCTCAATTATGCGGCAGATTGGCAGAGAACGCTTGGCGGTGTGTGAAACT	131
Qy	295	CAACCAACCCGAAGCCATTGCTTACATTAGTCCCATATTATGACGAAAGCGCGCTGG	354
Db	132	CAATTACACCAAGCGGTCCGCTCATTTAGCGGGGTGTGATGAAAGGCGCGTGTGG	191
Qy	355	AAAAAAACCCTGCCAGCTTTATGGAAGAGTGATGCACCTTTTTGAAAAAGATGAAGT	414
Db	192	TAATAAAAGCGTGGCGGATTTGATGCAAGAGGCGAGGACTTGGCTTAAAAAAGAAATGT	251
Qy	415	AATGCCCGGGGTGGTAAATAGTTCCGATCTAGGTGTAGAACCCCTTTCCTGATGG	474
Db	252	GATGGACGGCGTAGCAAGCATGATTTCATGAAGTGGGGATTGAAGCTAACTTCCCGATGG	311
Qy	475	TACGAAATCTGTAACCTGTGAATTGGCCCATCGAACCATGAGCACCTTCAAAGCGGGCA	534
Db	312	AACCAAGCTTGTAATCTCACTCCGGTAGAGTAATGGCAATTTAGCCCCCGGGCA	371
Qy	535	AGTGAATTTTGGTTGCGATAAAGACATCGAGCTCAATGCGAGCAAGAGTAAACCGAAT	594
Db	372	GGT---CTTCTTAAAAATGAGGACATTACTATTAAACGCGGCAAGAACCATTAGCTT	428
Qy	595	TGAGTTTACTAATGAAGGCTTAATCTTGCTGATGGGTAGGACATTTCCACTCTTTGA	654
Db	429	GAAAGTGAATAAAGCGCATCTGCTGTGTCAGGTGGGATCACATTTCCACTTTCTTGA	488
Qy	655	AGCTAACAGGCATAAATTCGATCGTGAAGAAAGCCATGCGCAAGCCCTAGATATTCC	714
Db	489	AGTGAATAAGCTTTGGACTTCGATCGCGCAAAAGCTTTTGCAACGCTTAGACATTGC	548
Qy	715	CTCTGGCAACACGCTAGCCATTGGGGGAGGACAAACCGCAAGGTGAGTTGATTCCTCT	774
Db	549	ATCTGGAACAGCGGTGGCTTTGAAACCGGGAGGAAAAAAGTGTGGNACTCATTTGACAT	608
Qy	775	TGTTGGCAGTAATAAAGTGTGTCATGAACGGGCTTTGTAATAACATCGCGAT---GA	831
Db	609	CGCGGGAATAAGCGCATCTATGCTTTAAATCTTGTGTGGATCGCAAGCCGATGCCGA	668
Qy	832	ACGCCATAAC-----ATAAGCGCTTGACAA-----GGGGAATCTCAGG	873
Db	669	TGTTAAAAAACTCGGCTTAAACCGCTTAAAGAAAAAGGTTTGGGTGTGTAAACTCGG	728
Qy	874	ATTT---ATCAAGTAAGGAGACTC-----CCATGAAATGA---AAAAACAAGA	916
Db	729	TTGTGAAGCACTAAGATTAACAAATGAAGNAACCAATGNAAGAATTTTCAAGNAAGA	788
Qy	917	ATATGTAATACCTACGGACCCACAAAGCGATTAAGTGGCTTAGGAGATACCGATCT	976
Db	789	ATATGTTTCTATGATGTTCCCACTACCGGGGATCGTGTGTAGACTCGCGGACATGATTT	848
Qy	977	TTGGGCGAAGTAGAATGACATATACCACTATGGCGAAGAACTTAAATTTGGCGGG	1036
Db	849	GATCTTAGAAGTGAGCATGATTGCAACCATTTATGGTGAAGAGATCAAAATTTGGGGCGG	908

Qy	1037	TAAAACTATCCGTGAGGGTATGSGTCAGAGCAATAGCCCTGATGAAAAACACCCTAGATTT	1096
Db	909	TAAAACTATCCGTGATGGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTT	968
Qy	1097	AGTCATCACTAACCGGATGATTTATCGACTACACCGGGATTTACAAAGCCGACATTTGGAT	1156
Db	969	GGTGTCTACTAAGCCCTCATTTGTGACTATAGGGCAATTTACAAAGCCGACATTTGGAT	1028
Qy	1157	TAAAAACGCAAAATCCATGGCAATGGCAAGGAGGAAACAAAGACATGCAAGATGGCGT	1216
Db	1029	TAAAGACGCAAGATTTGAGGCAATTTGGCAAGGAGGCAATTAAGACATGCAAGATGGCGT	1088
Qy	1217	AAGCCCTCATATGCTGTGGGTGGGCACAGAACCACTAGCAGGGGGAAGATGATTAT	1276
Db	1089	AGATAATATCTTTGCGTAGGTCCTGTACAGAGGCTTTGGCAGCTGAGGGCTTGATTGT	1148
Qy	1277	TACCCTCGGGGAATCGATTCACACACCACCTTCTCTTCTCCACAACTTCCTTACCGC	1336
Db	1149	AACCGCTGTGCAATCGATACGCATATTCACCTTATCTCTCTCCCAACAAATCCTTACTGC	1208
Qy	1337	TCTAGCCAAATGGCGTTACAAACATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGAA	1396
Db	1209	TTTTGCCAGCGGGTTACAAACCATGATTGGAGGAGCAGAGACCTCGGATGGCAGAA	1268
Qy	1397	TGCGACTACTATCACTCCGGGCAATGNACTTGCACCGCATGTTGGCGCAGCAGAA	1456
Db	1269	TGCGACCAACCATCACTCCCGGACGCGCTAATCTAAAGATGATTGTCGTGCAGCCGAA	1328
Qy	1457	GTATTTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTTAGCAAAAAACAACTTGT	1516
Db	1329	ATAACCCATGAATCTAGGCTTTTGGCTTAGGGGAATGTCTTAGCAACCTCTTTACG	1388
Qy	1517	AGAAACAGTAGAAGCGGGCGGATTTGGTTTAAATTTGCAATGAAGACTGGGGCAACACC	1576
Db	1389	CGATCAGATTGAAGCAGGCGGATTTGGTTTAAATTCACGAGACTGGGGAGCACACC	1448
Qy	1577	AAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATAGATGTCAGATTTGTATCCA	1636
Db	1449	TGCACTATTACCACTGCTCAATGTGCGCCATGAATACGATGTGCAAGTGTCTATCCA	1508
Qy	1637	CACGATACAGTCAATGAGGAGGTTATGTAGATGACACCTTAAATGCAATGAACGGCG	1696
Db	1509	CACGATACCTTAAAGAGGCGGCTGTGTAGAAGACACCTTAGAGGCGATTCGCGGGG	1568
Qy	1697	GGCAATCATGCTCACTACCAATTTAGGGAGCGGCTGTGTAGAAGACACCTTAGAGGCG	1756
Db	1569	CACCATCTACCTTCCACACTGAAGGGCTGGGGTGGACACGCTCCAGATGTTATCAA	1628
Qy	1757	CATGGCAGCGAGCTCAATATTTCTACCTCTCTCCACCCCACTATTTCCCTATACCAT	1816
Db	1629	AATGGCAGGGGAATTTAACTTCTACCGCTCTACTAACCGACCATTTCTTTCACCAA	1688
Qy	1817	TAAATCGTTGCAAGAACCTTAGACATGCTATGATGACATGCCACCACTAGACAAACGCAT	1876
Db	1689	AAACACTGAAGCCGAGCAGCATGGACATGTTAATGGTGTGCCACCACTTGGATAAAGAT	1748
Qy	1877	CGCGAGGATTTACAAATTTTCTCAAAAGCGTATCGCCCGGCTCTATCGCGCTCAAGA	1936
Db	1749	CAAGGAAGTGTGAGTTTGGGATTCGAGGATTCGCCCCCAACTATCGCGCTGAGAA	1808
Qy	1937	TGTGCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGG	1996
Db	1809	CCAACTCCATGACATGGGGATCTTTCTTCTATCACAGCTCCGACTCTCAGGCTATGGGACG	1868
Qy	1997	TGCGAGGCAATGATCTCTCGAACTTGGCAGACTCGGGATTAAGATTAAGAAAGATTTGG	2056
Db	1869	CGTAGGCGAGGTGATCAACGCACTTTGGCAGACAGCAGACAAAAACAAAAAGATTTGG	1928
Qy	2057	TAAAGTCTCTGAAGTGGCAAGATAAGATAATTTTCGCAATTTAGCGCTTACATCTCAA	2116
Db	1929	CGGCTTGAAGAGGAAAAAAGGCGATTAACGACACTTCCGCATCAACACGCTTACTCTAA	1988
Qy	2117	ATACACTATCAACCCCGCTTTTGACCCACCGCGGTGAGCGAGTATATCGGCTCTGTGAAGA	2176

```
Db 1989 ATACCATCAACCCCGGATCGGCATGGATTTCTGACTATGTGGCTCTGTGGAAGT 2048
Qy 2177 GGGCAAGATCCCGACTTGTGGTGTGGAATCTCTGCTCTTTTGGCGTAAACCCAAAT 2236
Db 2049 GGGCAATACCCGACCTCGTCTTGGAGTCGGCTTTCTTGGCATTAAGCCCAATAT 2108
Qy 2237 CGTGATCAAGGGCGGTATGTGGTCTTCTGAAATGGGCGATTCTAACGGCTGTGTC 2296
Db 2109 GATTATTAAGGGCGGATTTATTGGGCTCTCTCAAAATGGGCGATGCCATGGCTTATTC 2168
Qy 2297 CACTCCCAACCGGTTATTATCCGCGAAATCTTTGGGCATCACGSCAAGGCGAAATTGA 2356
Db 2169 CACCCTCAGCCCGCTATTATCCGCGAAATGTTTGGACACCATGGGAAAAACAATTCGA 2228
Qy 2357 CACCAGCATCACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAAAGCTGGG 2416
Db 2229 CACCAATATCACTTCGTGTCCTCAAGCGGCTTACAGCGCAGGATCAAGAAAGACTAGG 2288
Qy 2417 CTTAGAGCGCAAGTTCTTACCGGTCAAAAATCGCGTAACTACCAAGAAAGACTTCAA 2476
Db 2289 GCTAGATCGCGCGCACCGCAGTGAATACTGTGCAATATCACTAAAGAGGACCTCAA 2348
Qy 2477 GTTCAAGCAGAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGTCTTTGTAGA 2536
Db 2349 ATTCAACGATGTGACCGCACATATTGATGTCAACCCCTGAAACCTATAAGGTGAAGTGA 2408
Qy 2537 TGGCAAACTCTGCACTCTTAAACCCACCTCCGCAAGTGCCTTAGCCACAGCGCTACACTTT 2596
Db 2409 TGGCAAGAGTAACCTCTAAAGCAGCAGATGATTTAGCCCTAGCCCACTTTATATTT 2468
Qy 2597 CTCTAGG 2604
Db 2469 GTTCTAGG 2476
```

RESULT 4

```
US-08-432-697-19
; Sequence 19, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
```

```
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence."
US-08-432-697-19

Query Match 35.0%; Score 1008.8; DB 3; Length 2619;
Best Local Similarity 66.2%; Pred. No. 8.9e-241;
Matches 1634; Conservative 0; Mismatches 792; Indels 42; Gaps 8;

Qy 175 TGCCTTTAAATTTAAACAAGAGTAAATAGTGAACTCACACCCAAAGAGCAAGAAA 234
Db 13 TACCAATAGAAATTCATTAAGAGAGTTAGGATGAACCTAACGCTTAAAGAACTAGACAA 71
Qy 235 GTTCTTGTATTATATGCGGCGAAGTGGCTAGAAAGCCAAAGCAGAGGGCTTAAAGCT 294
Db 72 GTTAATGCTCCATATATGCGGCGAGATTTGGCAGAGAACCTTGGCGCGTGGTGTGAACT 131
Qy 295 CAACCAACCCGAAAGCCATTTGCTTACATTTAGTCCCATATATTATGGACGAAAGCGCGCTGG 354
Db 132 CAATTACACGAGCGGTGCGCTCATTTAGCGGCGGTGTGTGTAAGAGCGCGTGTG 191
Qy 355 AAAAAAACCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAAGAAAGATGAAGT 414
Db 192 TAATAAAGCGTGGCGGATTTGATGCAAGAGCAGGACTTGGCTTTAAAGAAAGAAATGT 251
Qy 415 AATGCCCGGGTGGTAAATATGTTCCCGATCTAGTGTAGAGCCACCTTTCTCTATGG 474
Db 252 GATGGACCGCGTAGCAAGCATGATTCATGAAGTGGGATTTGAAGCTAACTTCCCGATGG 311
Qy 475 TACGAAACTTGTAACTGTGAATTTGGCCCATCGAAACAGATGAGCACTTCAAAGCGGCGCA 534
Db 312 AACCAAGCTTGTAACTATCCACACTCCGCTAGAGGATATGCAAAATTAGCCCGCGCGA 371
Qy 535 AGTGAATTTGGTTCGATAAAGACATCGAGTCAATGAGGCAAGAGAGTAAACCGAACT 594
Db 372 GGT---CTTCTTAAAAAATGAGGACATTAATACTTAACGCGCGCAAGAGCACTTAGCTT 428
Qy 595 TGAGTTACTAATGAAGGCGCTAAATCTTGTGATGGGTAGGCATTTCCACTTCTTTGA 654
Db 429 GAAAGTGAATAAAGGCGATCGCTGTGAGGTGGGATCACATTTCCACTTCTTTCGA 488
Qy 655 AGCTAAACAGGCACCTAAATTCGATCGTGAAGAGCTATGGCAAGCGCTAGATATTCC 714
Db 489 AGTGAATAGCTCTTGGACTTCGATCGCGCAAAAGAGCTTTTGAAGAGCGCTAGACATTC 548
Qy 715 CTCTGGCAACACGCTACGCTATGGGGGAGGACAAACCCGCAAGTGCAGTTGATTCCTCT 774
Db 549 ATCTGNAACAGCGGTGCGCTTTTGAACCCGGGAGGAAAAAGTGTGGAACCTATTGACAT 608
Qy 775 TGGTGGCAGTAAAAAGTGAATTTGGCATGAACGGGCTTGTGTAATACATCGCGAT---GA 831
Db 609 CGCGGGAATAGCGCATCTATGGCTTTAATTTCTTTTGGTGGATCGCCAAAGCGCATGCCGA 668
Qy 832 ACGCCATAAAC-----ATAAGCGCTTTGACAA-----GGCGAAATCTCACGG 873
Db 669 TGGTAAAAAACTCGCGCTTAAACCGCGCTAAAGAAAGGTTTTTGGTCTGTAACTGCGG 728
Qy 874 ATTT--ATCAAGTAAGGAGACTC-----CCATGAAATGA---AAAAACAAGA 916
```



```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence..
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence..
; US-08-466-248-19

Query Match 35.0%; Score 1008.8; DB 3; Length 2619;
Best Local Similarity 66.2%; Pred. No. 8.9e-241;
Matches 1634; Conservative 0; Mismatches 792; Indels 42; Gaps 8;

Qy 175 TGCCTTTAAATTTAAACAAGGAGTAAATAGGTGAAACTCACACCAAGAGCAAGAAAA 234
Db 13 TACCAATAGAAATTCATAGAGGCTTTAGGATGAACTAACGCTTAAGAACTAGACAA 71

Qy 235 GTTCTTGTATATATATGCGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCT 294
Db 72 GTTAATGCTCCATATATGCGGGCAGATTGGCAGAGAAACGCTTGGCGCGTGTGAAACT 131

Qy 295 CAACCAACCGAAGCCATTGCTTATAGTCCCATATATTAGGAGAGCGCGCGGTGG 354
Db 132 CAATATACACCGAAGCGGTGCGGCTCATTAGCGGGCGTGTGATGTAAGAAAGCGCGTGG 191

Qy 355 AAAAAAACCGTTGCGAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGT 414
Db 192 TAATTAAGAGCGTGGCGGATTTGATGCAAGAGCGAGGACTTGGCTTAAAAAGAAATGT 251

Qy 415 AATGCCCGGGTGGTAAATATAGTTCCCGATCTAGGTGTAGAGCCACTTTCTCTGATGG 474
Db 252 GATGACCGCGGTAGCAAGCATGATTCATGAAGTGGGGATTAAGCTAACTTCCCGCGATGG 311

Qy 475 TACGAACTTGTAACTGTGAATGGCCCATCGAAGCAGATGAGCACTCAAGCGGGCGA 534
Db 312 AACCAAGCTTGTAACTATCCACACTCCGGTAGAGGATATAGGCAAAATAGCCCGCGGGA 371

Qy 535 AGTGAATTTGGTTCGATTAAGACATCGAGCTCAATGACAGCAAGAGTAACCGAATCT 594
Db 372 GGT----CTTCTTAAAAATGAGACATTAATAATACGCGCGCAAGAGCCATTAGCTTT 428

Qy 595 TGAGTTACTAATGAAGGCGCTTAAATCTTTCATGTGGGTAGCCATTTTCACTTCTTTGA 654
Db 175 TGCCTTTAAATTTAAACAAGGAGTAAATAGGTGAAACTCACACCAAGAGCAAGAAAA 234
Db 13 TACCAATAGAAATTCATAGAGGCTTTAGGATGAACTAACGCTTAAGAACTAGACAA 71
Db 72 GTTAATGCTCCATATATGCGGGCAGATTGGCAGAGAAACGCTTGGCGCGTGTGAAACT 131
Db 295 CAACCAACCGAAGCCATTGCTTATAGTCCCATATATTAGGAGAGCGCGCGGTGG 354
Db 132 CAATATACACCGAAGCGGTGCGGCTCATTAGCGGGCGTGTGATGTAAGAAAGCGCGTGG 191
Db 355 AAAAAAACCGTTGCGAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGT 414
Db 192 TAATTAAGAGCGTGGCGGATTTGATGCAAGAGCGAGGACTTGGCTTAAAAAGAAATGT 251
Db 415 AATGCCCGGGTGGTAAATATAGTTCCCGATCTAGGTGTAGAGCCACTTTCTCTGATGG 474
Db 252 GATGACCGCGGTAGCAAGCATGATTCATGAAGTGGGGATTAAGCTAACTTCCCGCGATGG 311
Db 475 TACGAACTTGTAACTGTGAATGGCCCATCGAAGCAGATGAGCACTCAAGCGGGCGA 534
Db 312 AACCAAGCTTGTAACTATCCACACTCCGGTAGAGGATATAGGCAAAATAGCCCGCGGGA 371
Db 535 AGTGAATTTGGTTCGATTAAGACATCGAGCTCAATGACAGCAAGAGTAACCGAATCT 594
Db 372 GGT----CTTCTTAAAAATGAGACATTAATAATACGCGCGCAAGAGCCATTAGCTTT 428
Db 595 TGAGTTACTAATGAAGGCGCTTAAATCTTTCATGTGGGTAGCCATTTTCACTTCTTTGA 654
```

```

Qy 1697 CGCATCATGCTTACCACTTGGAGGCGGTGGAGGACCTCCTGATGTTATCAC 1756
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1569 CACCATCATACCTTCCACACTGAAGGGGCTGGGGGTGGACCGCTCAGATGTTATCAA 1628
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1757 CATGGCAGGCGAGCTCAATATTCTTACCTCTCCACACCCCTACTATTCCTATACCAT 1816
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1629 AATGGCAGGGAATTAACATCTTACCCTCTTACTAACCAGCCTCTCTTCCACCA 1688
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1817 TAATACGGTTGACAGACATTTAGACATGCTCATGACATGCCACCACTAGACAAACGAT 1876
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1689 AAACACTGAAGCCGAGCAGATGGACATGTTAATGGTGTGCCACCTTGGATAAAGTAT 1748
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1877 CCGGAGGATTTACAAATTTCTTCAAGCCGTATCGCCCGGCTCTATCGCGGCTGAAGA 1936
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1749 CAAGGAAGATGTGCAGTTTGCAGTTTGCAGATTGCGCCCAAACTATCGCGGCTGAAGA 1808
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1937 TGTGCTCATGATATGGTGTGATCGGATGACAGCTCGGATTCGCAAGCAATGGGGCG 1996
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1809 CCAACTCATGACATGGGATCTTTCTTATCACCAGCTCCGACTCTCAGGCTATGGGACG 1868
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1997 TGCAGGCGAAGTGAATCTCTCGAACTTGGCAGACTCGGATAGAAATAAAGAAATTTGG 2056
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1869 CGTAGGCGAGTGATCACGCACTTGGCAGACGACAGACAAACAAAGAGTTGG 1928
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2057 TAAGCTTCTGAAGATGGCAAGATAACGATPAATTTCCGCAATTAAGCGCTACATCTCAA 2116
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1929 CGCTTGAAGAGGAAAGAGCGATAACGACAACTTCCGATCAAAAGCTACATCTCTAA 1988
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2117 ATACACTATCAACCCCGTTGACCCAGGCGGTGAGCGATATATCGCTCTGTGGAGA 2176
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1989 ATACACCATCAACCCCGGATCGCGCATGGGATTTCTGACTATGTGGGCTCTGTGGAAGT 2048
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2177 GGGCAAGATCGCGACTTGGTGTGTGGAATCTCGCTCTTTTTCGCGTAAACCCAAAT 2236
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2049 GGGCAATACCGCACTCGTGTGTGGAGTCCGGCTTTCTTGGCAATAGCCCAATAT 2108
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2237 CGTGATCAAGCGGGTATGGTGTCTTCTCTGAATGGCGGATTTCAACGCTCTGTGCC 2296
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2109 GATTATTAGGCGGATTTATGGCTCTCTCAATGGCGGATGCCATGCTATTC 2168
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2297 CACTCCCAACCGGTTATTACCGCGAAATTTTGGGATCAACGCAAGGCAATTTGA 2356
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2169 CACCCCTCAGCCGCTTATTACCGTGAATTTTGGACACCATGGGAAACAAATTCGA 2228
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2357 CACGAGATCACTTTGTTTCCAAAGTCGCTATGAATGGCGTGAAGAAAGCTGG 2416
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2229 CACCAATATCACTTGTGTCGCAAGCGGCTTCAAGCGGAGATCAAGAAAGAACTAGG 2288
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2417 CTTAGAGCGCAAGTTCTACCGGTCAAAACCTGCGCTAAACATCAACCAAGAAAGACTTCAA 2476
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2289 GCTAGATCGGCGGACCGCGAGTGAAGAACTGTGCAATATCAATAAAGGACCTCAA 2348
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2477 GTTCAACGACAAACGCAAAATTCACCGTGAATCCGAAACCTTCGAGGCTTTGTAGA 2536
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2349 ATTCAACGATGTACCGCACAATTTGATGTCAACCTTGAACCTTAAAGTGAAGTGA 2408
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2537 TGCAAACTCGCACTCTTAAACCCACTCGCAAGTGCCTCTAGCCAGCGCTACATTT 2596
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2409 TGGCAAGAGGTAACCTCTAAAGCAGCAGATGATTTGAGCCTTAGCGCACTTTATATTT 2468
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2597 CTTCTAGG 2604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2469 GTTCTAGG 2476
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 6

US-09-431-705-1
; Sequence 1, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia

```

; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (2)...(31)
; NAME/KEY: CDS
; LOCATION: (41)...(61)
; NAME/KEY: CDS
; LOCATION: (65)...(799)
; NAME/KEY: CDS
; LOCATION: (803)...(2512)
; NAME/KEY: CDS
; LOCATION: (2516)...(2692)
; NAME/KEY: CDS
; LOCATION: (2696)...(2896)
; NAME/KEY: CDS
; LOCATION: (2900)...(3322)
; NAME/KEY: CDS
; LOCATION: (3326)...(3385)
; NAME/KEY: CDS
; LOCATION: (3389)...(3406)
; NAME/KEY: CDS
; LOCATION: (3410)...(3466)
; NAME/KEY: CDS
; LOCATION: (3470)...(3598)
; NAME/KEY: CDS
; LOCATION: (3602)...(3661)
; NAME/KEY: CDS
; LOCATION: (3665)...(3697)
; NAME/KEY: CDS
; LOCATION: (3701)...(3769)
; NAME/KEY: CDS
; LOCATION: (3773)...(3817)
; NAME/KEY: CDS
; LOCATION: (3821)...(3844)
; NAME/KEY: CDS
; LOCATION: (3848)...(3889)

```

```

Query Match 34.9%; Score 1005.8; DB 3; Length 4824;
Best Local Similarity 65.8%; Pred. No. 6.6e-240;
Matches 1640; Conservative 0; Mismatches 812; Indels 39; Gaps 10;

Qy 150 AACGAAATTTAGGACTATATCCATTTGCTTTTAAATTTTAAACACAAG-GAGTAAATAGGTG 208
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 29 AAATGATCTGACGTACACAGCAATTTAGATATTAATCATCCACAGAGAGATCTCCATG 88
Qy 209 AAATCTCACCAACCAAGAGCAAGAAAGTTCTTTGTTATATTTATGCGGCGAAGTGGCTAGA 268
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 89 AAATCTCACCAACCAAGAGTTAGTAAAGTTGATGCTCCACTACGCTGGAGAAATTTGGCTAAA 148
Qy 269 AAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAGAGCCATTCCTTACATTAGTGCC 328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 149 AAACGCAAGAAAGAGGCAATTAAGCTTAATGTAGAAAGCAGTAGCTTTGATTAGTGCC 208
Qy 329 CATATTATGGACGAGCGCGCTGGAAGAAACCGTTGCGCAGCTTATGGAAGAGTGC 388
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 209 CATATTATGGAGAGAGGAGAGCTGGTAAAGACTGCGGCTGAATTTGATGCAAGAGG 268
Qy 389 ATGCACCTTTTGAAGAAAGATGAAGTAATGCCGCGGTGGGTAAATATGTTCCCGATCTA 448
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

US-09-431-705-1

Db 269 GCACCTCTTTTAAACACAGATGATGTGATGGATGGCGTGGCAAGCATGATCCATGAAGTG 328
Qy 449 GGTGTAGAACCACTTTCTTGATGGTACGAACACTTGTAACTGTGAAATTTGGCCCATCGAA 508
Db 329 GGTATGGAAGCGATGTTTTCTTGATGGGACTAAACTCGTAAACCGTGCAATACCCCTATTGAG 388
Qy 509 CCAGATGAGCACTTCAAGACGGCGGAAGTGAATTTGGTTGGGATAAAGACATCGAGCTC 568
Db 389 GCCAATGGTAAATTAGTTCTTGTTGGTGAGTTG---TTCTTTAAAAAATGAAGACATCACTATC 445
Qy 569 AATGCAAGCAAGAAAGTAAACCACTCTGAGGTGTACTAATGAAGGCCATAAATCCTTTGCAT 628
Db 446 AACCAAGCAAAAAGCCGTTAGCGTGAAGTGAATTAATAATGTTGGCAGACAGCCGTTCAA 505
Qy 629 GTGGGTAGCCAAATTCACATTTCTTTGAAGCTAACCAAGGCACATAAATTCGATCGTGA AAAA 688
Db 506 ATCGGCTCACACTTCCAATTTCTTTGAAAGTGAATAGATGCCTAGACTTTTGAACAGAGAAAA 565
Qy 589 GCCTATGCAACACGCTAGATATTCCTCTGCTGCGCACTAAAGTGCATGAAACGG 808
Db 566 ACTTTGCGGTAAACGCTTAGACAATTCGAGCGGGAAGCGGTAAAGATTTGAGCCTGGCGAA 625
Qy 749 ACCCGCAAGTGCAGTTGATTTCTCTGCTGCGCACTAAAGTGCATGAAACGG 808
Db 626 GAAAAATCCGTAGAAATTTGATGCAATTTGCGGTAAAGAGATCTTTGGATTTAACGCA 685
Qy 809 CTTGTGAATAACATCGCG---ATGAACGCCA-----TAAACA-----TAAAGCG 850
Db 686 TTGGTTGATACAGACAGACAAAGCAAGCAAAAAAATTTGCTTTACACAGAGCTAAAGAG 745
Qy 851 C-----TTGACAAGGG---AA---ATCTCAGGATTT---ATCAAGTAAAGAGACTCCC 896
Db 746 CGTGGTTTTCATGGCGCTTAAAGCGCATGACAACTATGTAAAGCAAAATTAAGAGGTAAAGAA 805
Qy 897 ATGAAA---ATGAAAAACAGAAATATCTAATACCTACGCAAGCCACCAAGGCGCATAAA 953
Db 806 ATGAAAAAGATTAGCAGAAAAGATATGTTCTATGTATGTCCTACTACAGGCGATAAA 865
Qy 954 GTGGCTTAGGAGATACCGATCTTTGGCGAAGAGTAGAACATGACTATACCACCTATGGC 1013
Db 866 GTGAGATTGGCGATACAGACTTGATCGCTGAGTGAACATGACTACACCATTTATGGC 925
Qy 1014 GAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGC 1073
Db 926 GAAGAGCTTAAATTCGTGGCGGTAAAAACCCCTAAGAGAAAGGCATGAGCCAATCTAACCAAC 985
Qy 1074 CCTGATGAACACACCTAGATTTAGTCACTACGCGATGATATGCACTACACCGG 1133
Db 986 CCTAGCAAAAGAGTTGGATTTAATTAATCACTAACGCTTTAATCGTGTGATTTACACCGGT 1045
Qy 1134 ATTTACAAAGCGCATTTGGGATTA AAAACCGCAAAATCCATGGCATTTGGCAAGGCAGGA 1193
Db 1046 ATTTATAAGCGGATTTGGTATTAAGATGGAATAATCGCTGGCATTTGGTAAAGCGGT 1105
Qy 1194 AACAAAGCATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTGTTGGGCACAGAAACA 1253
Db 1106 AACAAAGCATGCAAGATGGCGTTAAAAACAATCTTAGCGTAGGTCTGCTACTGAAAGCC 1165
Qy 1254 CTAGCAGGGGAAGGTATGATTAATTAACCGCTGGGGAAATCGAATTAACACCACTTCCTTT 1313
Db 1166 TTAGCCGTTGAAGGTTTGTATGTAACGCTGGTGGTATTGACACACATCCACTTCATTT 1225
Qy 1314 TCTCCAAACAATTCCTTACCGCTCTAGCCTAGGCGTTTACAACCATGTTTGGAGCGCGC 1373
Db 1226 TCACCCCAACAATCCCTTACAGCTTTTGCAAGCGGTGTAACACCATGATTTGGTGGTGA 1285
Qy 1374 ACAGTCTCTGTAGTGGACGAATCGCACTACTATCACTCCGGGCAAAATGGAACCTTGCAAC 1433
Db 1286 ACCGCTCTGCTGATGGCACTAAATCGCACTACTATCACTCCAGGCAGAGAAATTTAAAA 1345
Qy 1434 CGCATTTGCGGCGACAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAGGCAAT 1493
Db 1346 TGGATGCTCAGAGCGGCTGAAGAAATTTCTATGAATTTAGGTTTCTTGGCTTAAAGGTAAAC 1405

Qy 1494 AGCTCTAGCAAAAAACAACCTTGTAGAAACAAGTAGAAGCGGCGCGATTTGGTTTAAATTTG 1553
Db 1406 GCTTCTAAACGATGCGAGCTTAGCCGATCAAAATTTGAAGCGGCTGCGATTTGGCTTTGCAATT 1465
Qy 1554 CATGAAGACTGGGGCACACCAAGTGCATCGACTGCTTGTAGCGTGGCGAGAGTAA 1613
Db 1466 CAGGAAGACTGGGGCACCACTCTTCTGCAATCAATCATGCGTTAGATGTTGGGACAAA 1525
Qy 1614 TAGCATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGACAGGTTATGTAGATGAC 1673
Db 1526 TAGCATGTGCAAGTTCGCTATCGCCACAGACACTTTGAAATGAAGCCGTTGTTAGAGAC 1585
Qy 1674 ACCCTAAATGCAATGAAACGGGCGCGCATTCATGCTTACACATGAGGAGCGGTTGGA 1733
Db 1586 ACTATGGCTGCTATTTGCTGGAGCGCACTATGCACTTTCCACACTGAAGCGGCTGGCGC 1645
Qy 1734 GGACACTCACCTGATGTTATCACATGCGAGGCGAGCTCAATATTCTTACCTCTCCACC 1793
Db 1646 GGACACGCTCTGATATTATTAAGTAGCGGTGAACAACAATTTCTCCGCTTCCACT 1705
Qy 1794 ACCCCCACTATTCCCTATACCAATTAATACGTTGACAGAACACTTAGACATGCTCATGACA 1853
Db 1706 AACCACCACTCCCTTTACCGTGAATACAGAGCAGACACATGACATGCTTTATGGTG 1765
Qy 1854 TGCCACCACTTAGACAAAACGCAATCCGCGAGGATTTTACAATTTTCTCAAGCCGTATCCCG 1913
Db 1766 TGCCACCACTTGGAATAAAGACATTAAGAAGATGTTTCAAGTTGCTGATTTCAAGGATCCGC 1825
Qy 1914 CCGGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATCGGATGACAGCAAGC 1973
Db 1826 CCTCAACCAATTTGCGGCTGAAGACACTTTTGCAATGACATGGGATTTTCTCAATCACCACT 1885
Qy 1974 TCGGATTCGCAAGCAATGGGCGTGCAGCGAAGTGAATTCCTCGAACTTTGGCAGACTGCG 2033
Db 1886 TCTGACTCTCAAGCGATGGGCGGTGGGTGAAGTTTACTAGAACTTTGGCAACAGCT 1945
Qy 2034 GATAAGAAATAAAAAAGAAATTTGGTAAAGCTTCTCGAAGATGGCAAGATAACGATAATTC 2093
Db 1946 GACAAAAACAAGAAAGAAATTTGGCGCTTTGAAAGAAAGAAAGCGGATTAACGACAACTTC 2005
Qy 2094 CGCATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGGTGAGC 2153
Db 2006 AGGATCAACCGCTACTTGTCTAAATAACCATTAACCCAGCGATCGCTCATGGGATTAGC 2065
Qy 2154 GAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTGTGTGGAATCTCTGCC 2213
Db 2066 GAGTATGTAGGTTCAAGTAGAGTGGCAAGTGGCTGACTTGGTATTTGGAGTTCAGCA 2125
Qy 2214 TTTTGTGGCGTAAAAACCCCAAAATCGTGATCAAGGCGGTATGGTGGTCTTTCTCTGAAATG 2273
Db 2126 TTTTGTGGCGTAAACCCCAACATGATCATCAAGGCGGATTCATTGGTTAAGCCAAATG 2185
Qy 2274 GCGCATTTAAGCGCTCTGTGCCACTCCCGCAACCGGTTTATTACCGGCAATGTTTGGG 2333
Db 2186 GCGCATGCGAAACGCTTCTATCCCTACCCCAACACCGGTTTATTACAGAGAAATGTTGCGT 2245
Qy 2334 CATCAGCGCAGGCGAAATTTGACACCGACATCACTTTTGTTCCAAAGTCGCTATGAA 2393
Db 2246 CATCATGTGTAAGCTTAAATACGATGCAAAACATCACTTTTGTGTCTCAAGCGGCTTATGAC 2305
Qy 2394 AATGCGGTGAAGAAAGAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAACTGCGCT 2453
Db 2306 AAAGCATTTAAGAGAAATTTAGGACTTTGAAGACAGTGTTCGCGGTAAAAATTTGAGA 2365
Qy 2454 AACATCAACAGAAAGACTTTCAAGTTTCAACGCAAAACCGCAAAATCAACCGTGCATCCG 2513
Db 2366 AATATCACTTAAAAAAGACATGCAATTAACGACACACTACCGCTCAATTTGAAGTCAATCCT 2425
Qy 2514 AAAACCTTCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCACTTCGAGTG 2573
Db 2426 GAAACTTCCATGTGTTCGTGGATGGCAAAAGAAAGTAACCTTCTAAACGAGCCATAAAGTG 2485


```
QY 2574 CCTCTAGCCAGCGCTACACACTTCTCTAGG 2604
| | | | | | | | | | | | | | | | | |
Db 2486 AGCTGGCGCACTCTTTAGCAATTTCTAGG 2516
| | | | | | | | | | | | | | | | | |

RESULT 7
US-09-431-705-19
; Sequence 19 Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (3893)...(3934)
; NAME/KEY: CDS
; LOCATION: (3938)...(4027)
; NAME/KEY: CDS
; LOCATION: (4031)...(4285)
; NAME/KEY: CDS
; LOCATION: (4289)...(4300)
; NAME/KEY: CDS
; LOCATION: (4304)...(4408)
; NAME/KEY: CDS
; LOCATION: (4412)...(4471)
; NAME/KEY: CDS
; LOCATION: (4475)...(4598)
; NAME/KEY: CDS
; LOCATION: (4592)...(4669)
; NAME/KEY: CDS
; LOCATION: (4673)...(4711)
; NAME/KEY: CDS
; LOCATION: (4715)...(4774)
; NAME/KEY: CDS
; LOCATION: (4784)...(4824)
; US-09-431-705-19

Query Match 34.98; Score 1005.8; DB 3; Length 4824;
Best Local Similarity 65.88; Pred. No. 6.6e-240;
Matches 1640; Conservative 0; Mismatches 812; Indels 39; Gaps 10;

QY 150 AACGAATATTAGGACTAATATCCCAATTCCTTTTAAATTTTAAACACAAG-GAGTAATAGGTG 208
| | | | | | | | | | | | | | | | | |
Db 29 AATGATCTGAGCTACACAGCAATTTAGATATTAATCATCCACAGAGATCTCCATG 88
| | | | | | | | | | | | | | | | | |

QY 209 AAATCACCACCCAAAGAGCAAGAAAGTTCTTGTTATATTATTCGGCGCAAGTGGCTAGA 268
| | | | | | | | | | | | | | | | | |
Db 89 AAATCACCACCCAAAGAGTTAGATAGTTGATGCTCCACTACGCTGGAGAATTGGCTAAA 148
| | | | | | | | | | | | | | | | | |

QY 269 AACGGCAAGCAGGCTTAAAGCTCAACCAACCCGAGGCATTTGCTTACATTAGTGCC 328
| | | | | | | | | | | | | | | | | |
Db 149 AAACGCAAGAAAGGCAATTAAGCTTAACTATGTAGAAGCAGTAGCTTTGATTAGTGCC 208
| | | | | | | | | | | | | | | | | |

QY 329 CATATTATGAACGAGCGCCGCTGGGAAAAAACCCTTCGCCAGCTTATGGAAGATGC 388
| | | | | | | | | | | | | | | | | |
Db 209 CATATTATGAACGAGAGAGCTGGTAAAGAGCTCGGCTGATTTGATGCAAGAGGG 268
| | | | | | | | | | | | | | | | | |

QY 389 ATGCACCTTTTGAAGAAAGATGAAGTAATCCCGGGGTGGGTAAATATGTTTCCCGATCTA 448
| | | | | | | | | | | | | | | | | |
```


QY 1494 AGCTCTAGCAAAAACAACTTGTAGAACAGTAGAGCGGGCGGATGGTTTAAATTG 1553
DB |||||
QY 1406 GCTTCTAACGATCGAGCTAGCCGATCAAAATGAAGCCGGTGGGATTTGCAATT 1465
DB |||||
QY 1554 CATGAAGACTGGGCGCAACACCAAGTGCATCATCTGCTTGGAGCGGCAGATGAA 1613
DB |||||
QY 1466 CACGAAGACTGGGCGCACCACTCTTCTGCAATCAATCATGCTTGTAGATGTTGCCGACAAA 1525
DB |||||
QY 1614 TAGCATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGGTTATGTAGATGAC 1673
DB |||||
QY 1526 TAGCATGTGCAAGTCGCTATCGCCACACAGACTTTGAATGAAGCCGGTTGTGTAGAAGAC 1595
DB |||||
QY 1674 ACCCTAAATGCAATGAACAGCGGCGCCATCATGCTTACCAATTTCTGAGGGAGCGGGTGGGA 1733
DB |||||
QY 1586 ACTATGGCTGCTATTGTGTGAGCGCACTATGCACACTTTCCACACTGGAAGGCGCTGGCGGC 1645
DB |||||
QY 1734 GGACACCTACCTGATGTTATCACCATGCGAGCGCTCAATATTTCTACCCCTCCACAC 1793
DB |||||
QY 1646 GGACACGCTCTCTGATATTTAATAAGTAGCCGGTGAACAACTTTCTCCGCTTCCACT 1705
DB |||||
QY 1794 ACCCCCACTATTCCTCTATACCAATTAATACGGTGTGAGAACACTTAGACATGCTCATGACA 1853
DB |||||
QY 1706 AACCCCACTCTCTTCCACCGTGNATACAGAGCAGACATGGACATGCTTATGGTG 1765
DB |||||
QY 1854 TGCACCACTTAGACAAACCGCATCGGAGGATTTCAATTTTCTCAAAGCCGTTATCCGC 1913
DB |||||
QY 1766 TGCACCACTTTGATATAAAGCATTAAGAGATGTTCAAGTTCGCTGATTTCAAGGATCCGC 1825
DB |||||
QY 1914 CCGGCTCTATCGCGCTGAAGATGTGCTCATGATATGGTGTGATCGCATGACAGC 1973
DB |||||
QY 1826 CCTCAAAACCATTCGCGCTGAAGACACTTTGCATGACATGGGGATTTCTCAATCACCACT 1895
DB |||||
QY 1974 TCGATTCGCAAGCAATGGGCGTCGAGCGAAGTGTCTCGAACTTTGGCAGACTCGC 2033
DB |||||
QY 1886 TCTGACTCTCAGCGATGGGCGGTGGTGAAGTTATCATCTAGACTTGGCAACAGCT 1945
DB |||||
QY 2034 GATAAGCAATAAAAAGAAATTTGGTAAGCTTCTGAAAGTGGCAAGATAACGATAATTC 2093
DB |||||
QY 1946 GACAAAAACGAAGAATAATTTGGCGCTTTGAAAGAAAGAAAAAGGAGATAACGACAACTTC 2005
DB |||||
QY 2094 CGATTAAGCGCTACATCTCAATATACATATCAACCCCGCTTTGACCCACGGGTGAGC 2153
DB |||||
QY 2006 AGGATCAACCGCTACTTGTCTTAAATACACCACTTAACCCAGCGATCGCTCATGGGATAGC 2065
DB |||||
QY 2154 GAGTATATCGGCTCTGTGGAAGAGGGCAAGATCGCGACTTGGTGGTGAATCCTGCC 2213
DB |||||
QY 2066 GAGTATGTAGGTTCAAGTAAAGTGGGCAAGTGGCTGACTTGGTATTTGGAGTCCAGCA 2125
DB |||||
QY 2214 TTTTTCGGCGTAAACCCAAATCGTGATCAAAAGCGGTATGGTGTCTTCTGAAATG 2273
DB |||||
QY 2126 TTCTTTGGCGTGAACCAACATGATCATCAAGCGGATTCATTCGCTTAAGCCAAATG 2185
DB |||||
QY 2274 GGCATTTCAACCGCTCTGTGCCACTCCCAACCGGTTTATTAACCGGAAATGTTGGG 2333
DB |||||
QY 2186 GGCATGCGAAGCTTCTATCCCTACCCCAACCGGTTTATTAACAGAGAAATGTTGCT 2245
DB |||||
QY 2334 CATCAAGCGAGCGCAATTTGACACCAAGCATCACTTTGTTTCCAAAGTCGCTTACAA 2393
DB |||||
QY 2246 CATCATGTAAGCTAAATACGATGCAAAACATCACTTTTGTGTCTCAAGCGGCTTATGAC 2305
DB |||||
QY 2394 AATGGCGTGAAGAAAGAAAGCTGGGCTTAGAGCGCCAAAGTTTCTACCGGTCAAAAACCTGCCGT 2453
DB |||||
QY 2306 AAAGCATTAAGAAGAAATAGGACTTGAAGACAAAGTGTGGCGGTAAAAAATGTCAGA 2365
DB |||||
QY 2454 AACATCAACGAAGAAGCTTCAAGTTCAACGAACAAACCGGCAAAAATCACCGTCGATCCG 2513
DB |||||
QY 2366 AATATCACTAAAAAAGACATGCAATTTCAACGACACTACCGCTCACTTGAAGTCAATCCT 2425
DB |||||
QY 2514 AAAACCTTCGAGGCTTTGTAGATGGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTG 2573
DB |||||
QY 2426 GAAACTTACCCTATGTTGCTGGTGGCAAGAAAGTAAGTAACTTCTTAAACCAAGCAATAAAGTG 2485
DB |||||

QY 2574 CCTCTAGCCGAGCGCTACACTTTCTCTTAGG 2604
DB |||||
QY 2486 AGCTTGGCGCAACTCTTTAGCATTTTCTTAGG 2516
DB |||||

RESULT 8
US-07-732-242C-8
; Sequence 8, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeshi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frischauf, Holtz, Goodman & Woodward, P.C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; COMPUTER: IBM PC compatible (NEC PC-9801 RX)
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII Form
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,242C
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herbert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET NUMBER: 910532/HG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)972-1400
; TELEFAX: (212)370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6131 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-732-242C-8

Query Match 23.2%; Score 667.4; DB 2; Length 6131;
Best Local Similarity 59.0%; Pred. No. 1.3e-155;
Matches 1460; Conservative 0; Mismatches 951; Indels 63; Gaps 11;

QY 185 ATTTAACACAAAGGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAAGTTCTTTGTTA 244
DB |||||
QY 317 ACTTAAATAGGAGGTATCCATGAACCTGACTTCAGGTGAATGGAAGGCTCATGATT 376
DB |||||
QY 245 TATTATGCGGGCGAAGTGGCTAGAAAGCGCAAGAGAGGGCTTAAAGCTCAACCAACCC 304
DB |||||
QY 377 GTAGTGGCGGCTGACTTGGCCCGCTCGTAAAGAGCGGGGCTTAAAAATTATATCTCT 436
DB |||||
QY 305 GBAGCCATGCTTACATTTAGTCCCATATATTGAGAGAGCGCGCTGGAAGAAAAAAC 364
DB |||||
QY 437 GAAGCTGCGCAATGATTCATATGAAAGTCTGGAGGGGCGGGATGG---AAAAACG 493
DB |||||
QY 365 GTTGCCCGAGCTTATGGAAGAGTGCATGCACCTTTTCAAAAAAGATGAAGTAATGCCCGGG 424
DB |||||
QY 494 GTAGCTGAGTTAAUGCAATACCGTGCACAGGATTTCTACAAAGAGAGATGTAATGGAGGG 553
DB |||||
QY 425 GTGGGTAATATGTTCCCGATCTAGGTGTAGAGGCCACTTTCCTGATGTTAGAAACTTT 484
DB |||||

Db 554 GTTCCGGAATGATCCCGGATATTTCAAATTGAGGCAACCTTTTCTCTGATGGAAACAAAGCTT 613
Qy 485 GTAACCTGTAATTTGGCCCATC--GAACACAGATGAGACTCTCAAAGCGGGCGAAGTGAAT 542
Db 614 GTACCGTTTCATGACCCGATCCGTTAAATGAGGAGGACGTACGATGATACACAGGGAGTA 673
Qy 543 TTGGTTCCGATAAAGAC--ATCGAGCTCAATGACAGGCAAAAGAACTAACCAGCACTTGAGGT 600
Db 674 TGTATTAAAAAAGAACCTATTTTATGCAATCAAAATAAGCAGACGATCAAGATTCGCGT 733
Qy 601 TACTAATGAAGGGGCTTAATCTTTGCAATGTGGGTAGCCATTTCCACTCTTTTGAAGCTAA 660
Db 734 GTTAAACCGGGCGATCGACTGTTTCAAGTTGGTTCCCATTTTCATTTTGAAGTAA 793
Qy 661 CAAGGCACTAAATTCGATCGTGAAGAAAGCTATGGAAGAACCCCTAGATATTCCTCTGG 720
Db 794 TCAATCGCTTCAATTTTCATCGTGAAGAAAGCATTTGGCATGCTTTTGAATATTCGGGCTG 853
Qy 721 CAACCGCTACGATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTGGTGG 780
Db 854 AACGGCGGTTTCGCTTCGAGCCCGAGATGCGAAAGAGTAGAAATAATTCATTTTCAGG 913
Qy 781 CAGTAAAAAAGTAGTTCGCGATGACGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
Db 914 TGAACGCAAGTGTATGG-----TTTAAATATGTAACGATG-----951
Qy 841 ACATAAGCGCTTGACAAAGCGGAATCTACGGATTTATCAAAGTAAGGAGACTCCCATGA 900
Db 952 -----GATCAGTTTGAATGGGAAAAAGAAATGAGTT-----TTTCGATGTC-995
Qy 901 AAATGAAMAAACGAATATGTAAATACCTACGACCCCAAGGCGATAAAGTGGCT 960
Db 996 -----GAA-----AGCAATATGCGGATATGTTTGGACCAACTGTGCGGCGACCCATTCGTT 1046
Qy 961 TAGGATACCGATCTTTGGGCGAGAGTAGAATGACTATACCACTATGCGGCAAGAC 1020
Db 1047 TGCAGATTCAGAAATGTTTATCGAAATGAAAGGAGCTATACAACTGTATGGAGATGAGG 1106
Qy 1021 TTAATTTTGGCGGGTAAACCTATCCGTGAGGATGAGGTCAG-AGCAATAGCC--CTG 1077
Db 1107 TAAAGTTTGGCGGGCGAAGGTGATCCGAGATGGAATGGGGCAGCATCTTTGGGACAA 1166
Qy 1078 ATGAAACACCCCTAGATTTAGTCACTAACCGGATGATTAATGCACTACACCGGATTT 1137
Db 1167 GCGATGAATGCGTCCGATCTCGTATTAAACAAATGCGATTAATGTTGAATACACAGTATTT 1226
Qy 1138 ACAAAGCGACATTTGGATTTAAAAACGGCAAAATCCATGCAATGCGAAGCGAGGAACA 1197
Db 1227 ATAAAGCAGATATCGGCATAAAGATGGAATGATTCCTCCATAGGAAAGCGGGGAACC 1286
Qy 1198 AGGACATGCAAGATGGCGTAAGCCCTCATATGTTGTTGGGTGGGACAGAACACTAG 1257
Db 1287 CGTTGTTAATGGACGGGT-----CGATATGTTGATTTGAGCGAGCAACGAAGTCATAG 1340
Qy 1258 CAGGGGAAGGTATGATTTATPACCGCTGGGGGAATCGATTCACACACCCACTTCCTTTCTC 1317
Db 1341 CCGCAGAAGGGATGATTTGTGACAGCGGAGGAATAGATGCTCATATTCATTTATTTGCC 1400
Qy 1318 CACAAATTCCTACCGCTCTAGCCATGCGGTTACAAACATGTTTGGAGCGGCGACAG 1377
Db 1401 CTCAGCAAAATCGAAACCGCTCTTTGCAATCGGGTGTGACCACTATGATTTGGGAGGAACAG 1460
Qy 1378 GTCCTGTAGATGACGATGCGACTACTATCACTCGGGCAAAATGGAATTCGACCGCA 1437
Db 1461 GACCCGCTACAGGCACAAAATGCGCACTTTGTACCGGGGCGCTTGGAAATATCATCGTA 1520
Qy 1438 TGTTCGCGCAGCAGGAAGAGTATTTCTATGAATGTGGGCTTTTTTGGGCAAAAGGCAATAGCT 1497
Db 1521 TGCCTCAAGCAGCGGAAGAAATTCGCGTAACTTTGGGCTTTTTTAGGAAAGGGAACCTGTT 1580
Qy 1498 CTAGCAAAAAACAACTTTGTAGAAACAGATGAGAGCGGCGCGATGTTGTTTAAATTCGATG 1557
Db 1581 CAGATGAGGCTCTTTTAAAGGAACAAATTTGAAGCGGAGCGGTGGGATTAAGCTTCACG 1640

Qy 1558 AAGACTGGGGCACACACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATACG 1617
Db 1641 AAGATTGGGGATCGACGGCGCGCTATTGATACATGTTTGAAGTGGCGGATCGATATG 1700
Qy 1618 ATGTCAAAGTTTGTATTCACACCGGATACAGTCAATGAGGACAGGTATGTAGATGACACCC 1677
Db 1701 ATGTGCAAGTAGCGATTCATACAGACATTTAAATGAGGCGGATTTGTTCGAGTACTTT 1760
Qy 1678 TAAATGCAATGAACGGGCGCCATTCATGCTCTACACATTCAGGAGCGGGTGGAGGAC 1737
Db 1761 TGAAGCCCATAGACGGTGCAGTGAATTCATACCTATCATACAGAAAGGGCTGGCGGGGAC 1820
Qy 1738 ACTACCTGATGTTATTCACCATGCGGAGGAGCTCAATATTCACCTCTCCACACCC 1797
Db 1821 ATGCTCCGGATATTATAAAGCGGCGGCTTCCCGAATATTTTGGCTTCTTCACGAATC 1880
Qy 1798 CCACATTTTCCCTATACCATTAATACGTTTGAGAAACACATTTAGACATGCTCATCACATGCC 1857
Db 1881 CAATCGACCTTATACTATCAATATCTTTGGAAGAGCATTTAGATATGTTAATGGTTTGGC 1940
Qy 1858 ACCACTAGACAAACGCGATCCGCGAGGATTTACAAATTTTCTCAAAAGCCGTATCCGCCCG 1917
Db 1941 ACCACTAGACGCTAATATTCAGAGGATATTTGCTTTTTCGCGATTCACGCATACGGAAG 2000
Qy 1918 GCTCTATCCGGCTGAAGATGCTCCATGATATGCGGTGTGATCGCGATGACAAGCTCGG 1977
Db 2001 AGACCATCCGCGCGAAGATGTTTACATGATTTAGGGCGTTTTTCAGCATGATTTTCGTCG 2060
Qy 1978 ATTCGCAAGCAATGGGCGTGCAGCGAAGTATTCCTCGAATTTGGGACACTTGGCGGATA 2037
Db 2061 ATTCACAGCGATGGGCGGAGTAGGAGAGTGAATCATTCGTACGTGGCAACGGCTGACA 2120
Qy 2038 AGAATAAAAAAGAAATTTGTTAAGCTTCTCTGAAGATGGCAAGATAACGATAATTTCCGCA 2097
Db 2121 AGATGAAAAGCAAGAGAGGAGTTTACAGAGACAATGGTGTGGGAGACAATTTTCGTG 2180
Qy 2098 TTAAGCGCTACATCTTCCAAATACATATCAACCCCGCTTTGACCCACGCGGTGAGCGAGT 2157
Db 2181 TGAAGCGTTATTTGCGCAATATACGATCAATCCGCGCATTCGTCTATGTTATTCGGAT 2240
Qy 2158 ATATCGGCTCTGTGGAAGAGGCAAGATCGCGCATCTGGTGTGTGGAATCTTGCCTTTT 2217
Db 2241 ATGTGGGTTCTGTTGAAGTGGGAAATTTAGCTGATTTAGTGTGTGGAATCTTGCCTTTT 2300
Qy 2218 TTGGCGTAAACCCCAAAATCGTGAATCAAAAGCGGTATGTTGCTTCTCTGAAATCGGCG 2277
Db 2301 TTGGTGTGAACCTGACTGGTCTTTAAAGGAGGAATGATTCCTTACGACACTATGGAG 2360
Qy 2278 ATTCTTAACGCGCTCTGTGCCCATCTCCCAACCGGTTTTTATTTACCGGAAATTTTGGGCATC 2337
Db 2361 ATCCCAATGCCAGCATTCGAGACCGGAGCGGTTTTTATATCGTCCGATGTTTGCAGCGA 2420
Qy 2338 ACGGCAAGCGGAAATTTGACACCGAGATCACTTTTGTGTTTCCAAAGTCGCTATGAAAAATG 2397
Db 2421 AAGGAGATGCCAAATATCAAAACGTCATACACCTTTGTTTTCGAAAGCAGCGTATGAAAAAG 2480
Qy 2398 GCGTGAAGAAAGCTGGGCTTTAGAGCCCAAGTTCTACCGGTCAAAACCTCGCGTAACA 2457
Db 2481 GCATTCATGAACAGTTGGGTTTGAAGAAAAAGGTGAACCAAGTCCATCGAATTCGAAAT 2540
Qy 2458 TCACCAAGAAAGACTTCAAGTTTCAACGACAAAAACCGCAAAAAATCACCGTCGATCCGAAAA 2517
Db 2541 TGACGAAAAAAGATTAATTTTGAACGATAAAAACCCCAAAATTTGACGTCGATCTCTCAGA 2600
Qy 2518 CTTTCGAGGCTTTTGTAGATGCGAAACCTCTGCACCTCTAAACCCACCTCGCAAGTGCCTC 2577
Db 2601 CATATGAAGTAAAGGTAGACGCTCAATTTAGTGACATGTGAACCCGCGCAGAAAACTGTCCTTA 2660
Qy 2578 TAGCCGAGCGCTACACTTTTCTTAGGCACAATGCCCCCTTTTGGGGGAGGTTATTTTAG 2637
Db 2661 TGGCACACCGGATTTTCTTATTTTGGGTGAGAAAAACATGATGTTGAAAAAGTAGTCG 2720

Db	564878	AAATATGCACTTAACCTCCAGAGAACAGAAAAAAGTAACTGATGCTTTTCTCCGCGGCGAACT	564819
Qy	262	GGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGCAAGCCATGCTTACAT	321
Db	564818	TGCGGCAAAACGCAAGCAGCGCGGTAAATTAACCTATCCAGAAAATTAATGCTTATAT	564759
Qy	322	TAGTGGCCATATATGACGAGCGCGCGGTGGAAGAAAAAAGCCGTTGCCAGCTTATGGA	381
Db	564758	TGCTAGTCAATTTACAAGAGGCGCAGAGAGAGGAA---TGAGTGTAGCGGAAGTCATGCA	564702
Qy	382	AGAGTGCATGCACCTTTTGAAGAAAAAGATGAAGTAAATGCCCGGGGTGGGTAAATATGTTCC	441
Db	564701	ATATGGCGCAACACTTTTAACCGTTGATGATGTCATGGAAGGTGGCGGAAATGGTTCA	564642
Qy	442	CGATCTAGGTGTAGAAGCACCTTTCTCTGATGGTACGAAACCTGTGAACCTGTGACC	501
Db	564641	TGAAGTCCAGATTGAAGCTACTTTCCCGGATGGCAGCAAACTTGTTACCGTGCATAATCC	564582
Qy	502	CATC-GA-ACCAGATGACACTTCAAGCGGGCGAAGTGAATTTGG--TTGCGATAAAG	557
Db	564581	AATCAGATAAACCGWAGGGTGGGCTTTAGCCCAACAAAATAAAAATATCAATGTTGGGCTAA	564522
Qy	558	ACATCGAGCTCAATGCGAGCAAGAGTAACCGAACTTTGAGGTTACTAATGAAGGCGCTA	617
Db	564521	AGCCACCTTACAAGGACAAAGATATGATCCAGGCGAATACCAATTAGCGGAAGCGCA	564462
Qy	618	AATCCTTTCG--ATGTGGGTAGCCATTTTCCACTTTCTTTGAAGCTAACAAAGGCACATAAAAT	675
Db	564461	TATTTCTCGTAAATGTGCGCAGAAAAACGTAATAATCGAAGTACAAATTCAGGCGACCG	564402
Qy	676	CGATCGTGAAGAAAGCT---ATGGCAAGCGCTAGATATTCCTCTGCGCAACAGCTA--	730
Db	564401	CCCAATTCAGATTGGCTGCGCATTAACCAATTTTGTGAACCAATAATGATGCCCTTAATTTGA	564342
Qy	731	-CGCATTTGGGCGAGACAAACCCGCAAGTCAGTTGATTCCTCTTGGTGGCAGTAAAAA	789
Db	564341	CCGCACTTTGGCAGCTGGAATGCGCTTAAATGTTCCATCTGCGAATGCGGTGCTTTGA	564282
Qy	790	AGTGATTGGCATGAACGGGCTTGTGAATAACATCGCGATGAACGCCATAAACATAAAGC	849
Db	564281	ACCGGTGAAGTGAATCAGTGGAAATGATGTTGCTTTTGGTGGTAACCA-ATCATTTATG	564223
Qy	850	GCTTGACAAGCGGAAATCTCAGGATTTATCAAGTAAG--AGACTCCCATGAAATGAA	907
Db	564222	GTTTCCATAATCAAAATGATGCGAAATTTAATAGGTAGGCGAAGTGGCAATTAACAATTC	564163
Qy	908	AAAAACAAGATATGTAATACCTACGAGACCAACCAAGCGATAAAGTGGCTTAGGAGA	967
Db	564162	AAGAGCGCAATATGTAGCAACTTATGTTGCCAACAGTTTGGCGATAAAGTCCGTTTAGGCGA	564103
Qy	968	TACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTTATGCGGAAGCACTTAAATTT	1027
Db	564102	TACCAATTTATGGGCAACCATTTGAACAAGATTTATTGACCAAGGTGATGATGTAATTT	564043
Qy	1028	TGGCGCGGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAACAC	1087
Db	564042	TGGTGGCGGTAAAGCGTGGTGTGATGCTCAAGCGGTACGGCAACTCGCGACAA	563983
Qy	1088	CC-----TAGATTTAGTCATCACTAAACCGGATGATATGGA-CTACA-CCGGGAT	1135
Db	563982	TCCGAATGTAATGGATTTTGTGATTTACCAACGTTGATGATCATTTGATGTAAATTAGGCAT	563923
Qy	1136	TTACAAAGCCACATTTGGGATTTAAACCGCAAAATCCATGGCATTTGCAAGGAGGAA	1195
Db	563922	TATCAAGCCCATATTTGATTTCTGTATTTGGGCGTATTTGGGTATTTGGACAGCAGGTAA	563863
Qy	1196	CAAGCATGCAAGATGGCGCTTCAATATGTTGTTGGGTGGGTGGGCAAGAGCACT	1255
Db	563862	CCCTGACACCATGATTAACGTCACACCAAAATATGATTTATCGGTGAGCAGCAGGATTTCA	563803
Qy	1256	AGCAGGGAAGGTATGATTTATCCGCTGGGGGAATGCAATTCACACCACTTCCTTTC	1315
Db	563802	TAAACGTTGCAATTTAAATTTGAACCGCTGGTGGTATCGATACCCACATTCATTTATTTG	563743
Qy	1316	TCACAACAATTTCCCTAFCGCTCTAGCCCAATGGCGTTACAACCATGTTTGGAGCGGCAC	1375
Db	563742	TCACAACAAGCACCAACATGCAATTTGAAAGTGGCTTACCACTTAATTTGGTGGTGAAC	563683
Qy	1376	AGGTCTCTAGATGGCAGGAATGCGACTACTATCACTCCGGCAAAATGGAACCTTGACCG	1435
Db	563682	TGGCCCTCTGATGTTGATACACGCAACCACTTGTATCCCTGGCGCATGTTATGGAACG	563623
Qy	1436	CATGTTGGCGCGCAGCAGAGAGTATTTATGAATGTGGGCTTTTGGGCAAGCGCAATAG	1495
Db	563622	TATGTTTCAAGCGGCAGAACCTTTCGCGTAAACGTCGATTTTGGTAAAGGCAACTG	563563
Qy	1496	CTCTAGCAAAAAACAACCTTGTAGAACAGTAGAAGCGGCGGATGTTTAAAATTGCA	1555
Db	563562	TTCAACCCCTAGATCTCTCGCTGAGCAAAATGGAAGCGGTGCAATAGGTTTAAAAATCCA	563503
Qy	1556	TGAAGACTGGGCGACAAACCAAGTGGCATCACTCACTCTTTCAGCGTGGCAGATGAATA	1615
Db	563502	CGAAGACTGGGGTGCACAGCCCTGCGGTGATTAATCTGCTTTAAAAAGTAGCAGATGAAT	563443
Qy	1616	CGATGTGCAAGTTTCTATCCACACGATACAGTCAATGAGGCGAGGTTATGTAGATGACAC	1675
Db	563442	GGATATTCAAGTGGCCATTTCAACAGACACGCTTAATGAAGTGGCTTTTGGAGACAC	563383
Qy	1676	CCTAAATGCAATGAACGGGCGCCCATCCATGCCCTACCACTTTAGAGGAGCGGTGGAGG	1735
Db	563382	GATGAAGCGATTTGATGAGCAGTCAATCATCTTTCCATACGAGGGGCGAGTGGTGG	563323
Qy	1736	ACACTCACTGATGTTATCACCATGCGAGGCGAGCTCAATATTCACCTCTCCACCAC	1795
Db	563322	TCATGCACCTGACATCATTTAAAGCAGCGATGATTTCAAAAGTATACCTGCTTCAACCAA	563263
Qy	1796	CCCCACTATTCCTTACCATTAATACGGTTCAGAACTTAGACATGCTCATGACATG	1855
Db	563262	CCGACTGCTCGTTTACCAAAAACCAATTTGATGAACATTTGGATATGTTGATGTTG	563203
Qy	1856	CCACACCTTAGACAAACGATCCGCGAGGATTTACAATTTTCTCAAAGCGGTATCCGCC	1915
Db	563202	CCATCACTTAGATAAACCGCTGCCGAGACGCTAGCTTTTCCGATAGCCGTATCCGCC	563143
Qy	1916	CGGCTCTATCCGGCTGAAGATGCTCATGATATGGGTGATCGCATGACAGCTC	1975
Db	563142	TGAACCACTATGACAGCAGAGATTTTTCATGATATGGCGCTTCTCTCCATTTAGTTTC	563083
Qy	1976	GGATTCGCAAGCAATGGGCGCTGAGGCAAGTATTCCTCGAACTTTGGCAGACTCGGA	2035
Db	563082	AGACTCTCAAGCGATGGGACGTTTGGCGAAGTCGTTATTCGTACATGGCAAACTGAGA	563023
Qy	2036	TAAAGATAAAAAAGAAATTTGGTAAAGCTTCTGGAAGATGGCAAGATAAACGATTTCCG	2095
Db	563022	TAAAGATAAAATGCAACGCTGGTGAAGCT---AGGTAATGAAGGAAACGATAACTTCCG	562969
Qy	2096	CATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGCGTAGCGGA	2155
Db	562968	TATTAACCGATATATCGGAAATACACCATCAACCCAGCAATTCACATGGTATTGCGGA	562909
Qy	2156	GTATATCGGCTCTGTGGAGAGGCAAGATCCGCACTTGTGGTGTGGGAATCTGCTT	2215
Db	562908	GCATATGCTCGTTAGAAATGGGTAAATTCGAGATATCGTGTATGGAACCGATGTT	562849
Qy	2216	TTTTGGCGTAAAAACCAAAATCGTATCAAAAGCGGTATGGTGTCTTCTCTGAAATGGG	2275
Db	562848	CTTTGGCGTAAAACTGAAGTGTGATTAATAAGGCTTTATAGCTATCGGAAATGGG	562789
Qy	2276	CGATTTCAACGCTGTGTCCTCCCAACCGGTTTATACCGGAAATGTTTGGGCA	2335
Db	562788	CGATCCAAATGCTCAATTTCCAACACCGCAACCTGATTTCTACCGTCCCAATGACGTTG	562729
Qy	2336	TCACGGCAAGGCAATTTGACACGACATCACTTTTGTGTTTCCAAAGTCGCCTATGAAA	2395
Db	562728	ACAAGGCTTAGCAACCGCAACCAACAGCAGTATTTCTTTGTTTCAAGCGCGCTGAAAAAGC	562669

Qy	2396	TGGCGTGAAGAAAGCTGGGCTTAGCGGCCAAGTCTTACCGGTCAAAAACCTGCGGTAA	2455
Db	562668	TGATATTCTGTCGGAAGTTTCGGTTTTACAAAGAAACCAATGCTGTGAAGGCTGCGCAA	562609
Qy	2456	CATCACCAAGAAAGACTTCAAGTTCAACGACAAAACGCAAAAATCACGTCGATCCGAA	2515
Db	562608	CGTAGGTAAAAAGATCTGGTTTCATAATGATGTAAACCAACAATCTCTGTTGATGCTGA	562549
Qy	2516	AACCTCTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCACCTCGCAAGTGCC	2575
Db	562548	ACGTTATGAAGTTCGAGTGGACGGAGAGTTAATTACCTGTGAACCACTGGATAGGTACC	562489
Qy	2576	TCTAGCCGAGCGCTACACTTCTTCTTA	2602
Db	562488	ATTGGGTGACGGATATTTTCTTATTTCTA	562462

RESULT 11

```

US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g

```

```
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

Query Match      23.1%; Score 665.4; DB 3; Length 1830121;
Best Local Similarity 57.0%; Pred. No. 5.9e-154;
Matches 1452; Conservative 1; Mismatches 1057; Indels 37; Gaps 14;

QY      83 TTTTAAATAATTACTTATATCATATATAATAATAATATTACTTATATTAATAAGTAAAT 142
DB      564998 TATATAAGGAGGATTTTTTTCTTCAAGGGGCTTACAAAAAATCTGACAAATATTTT 564939

QY      143 AAAAAGTAACGAAATAGGACTATATATCCATTCGCTTTAAATTTAAACACAGGAGT-A 201
DB      564938 TCCTATTTTAGATTTTACGTAAATTCGCTGACTTTCGTATATAAATAAGGATGA 564879

QY      202 ATAGGTGAACTCACCCAAAGAGAGAAAAGTTCTTGTATTATATATATGCGGCGGAAAGT 261
DB      564878 AAATATGCACCTTAACTTCCAGAGAACACAGAAAAAATCGATGCTTTTCTCGCGGCGGAACT 564819

QY      262 GGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGGAGCCATTTGCTTACAT 321
DB      564818 TCGCGCAAAACGCAAGCAGCGCGCTTAAATTTAAACTATCCAGAAACTATTGCTTTATAT 564759

QY      322 TAGTGCCCATATATTGGACGAAGCGCGTGGAAAAAACAACCGTTGCCAGCTTATGGA 381
DB      564758 TGCTAGTCATTTACAAGAGGCGAGCAAGAGAAGAA---TGAGTGTAGCGGAAGTCATGCA 564702

QY      382 AGAGTGCATGCACTTTTGTAAAAAAGATGAAGTAATGCCCGGGTGGTGAATATGTTCC 441
DB      564701 ATATGGCGCAACACTTTTAAACCGTTGATGATGTCATGGAAGGTGTGGCGGAAATGTTCA 564642

QY      442 CGATCTAGGTGTAGAAAGCCACCTTTCCTGATCGTAGCAAACTTGTAACTGTGAAATGGCC 501
DB      564641 TGAAGTCCAGATTGAAGCTACTTTTCCCGATGGCAGAACTTGTTCACCGTCATATATCC 564582

QY      502 CATC-GA-ACCAGATGAGCACCTTCAAAGCGGCGGAAGTCAAAATTTGG---TTGCGATAAAG 557
DB      564581 AATCAGATAACCGWAGGGTGGGCTTTAGCCCAAAATAAATAATATCAATGTGTGGGCTAA 564522

QY      558 ACATCGAGCTCAATGCAGGCAAAAGAAAGTAACCGAACTTGAGGTTACTAATGAAGGCGCTTA 617
DB      564521 AGCCCAACCTTACAAGGAACAAAGATATGATCCAGGCGAATACCAATTAGCCGAGGCGA 564462

QY      618 AATCCTTGC--ATGTGGGTAGGCATTTCCACTTTTGAAGCTTAACAAGGCACATAAATTT 675
DB      564461 TATTCTCGCTAATGTGCGCAGAAAAACCGTAAAAATCGAAGTAACAAATTTCAAGCGCACCG 564402

QY      676 CGATCGTGAARAAGCCT---ATGGCAAAACGCTAGATATTTCCCTCTGGCAACACGCTA-- 730
DB      564401 CCCAATTCAGTTGGCTCGCATTAACAATTTTGAACCAATAATGCCCTTAAATTTGA 564342

QY      731 -CGCATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTTCCTTTGTTGGCAGTAAAAA 789
DB      564341 CGGCATTTTGGCACGTGGAAATGGCTTAAATGTTCCATCTGGCAATGCGGTGCTTTTGA 564282

QY      790 AGTGATTTGGCATGAACGGGCTTTGTAATAACATCGCGGATGAACGCCATTAACAATAAGC 849
DB      564281 ACCCGTGAAGTGAATCAGTGGAAATAGTTGCTTTTGGTGGTAAACA-AATCATTTATG 564223

QY      850 GCTTGACAAGGCGAAATCTCACGGATTTATCAAGTAAGG--AGACTCCCATGAATAAGAA 907
DB      564222 GTTTCATAATCAAAATTTGATGGCAAAATATAAGGTAGGCAAGATGGCATTAACAATTTTC 564163

QY      908 AAAACAAGAAATATGTAATAACCTACGACCCCAAAAGGCGATAAAGTGCCTTTAGGAGA 967
DB      564162 AAGAGGCAATATGTAGCAACTTATGTTCCACAGTTGGCGATAAAGTCCGTTTAGGCGA 564103

QY      968 TACCGATCTTTGGGCGAGAAAGTAGAACAATGACTATACCACTATGCGGAGAAACTTAAATTT 1027
DB      564102 TACCAATTTATGGGCAACCACTTTGAACAAGATTTATTGACAAAGGTGATGAGTGTAAATTT 564043

QY      1028 TGGCGGGGTAAAACTATCCGTGAGGGTATGGGTACAGCAATAGCCCTGATGAAACAC 1087
DB      564042 TGGTGGCGGTAAAGCGTCGCTGATGGTCTCAAGCGGTACGGCAACTCCGCGACAA 563983
```


Qy	1088	CC-----TAGATTTAGTCATCATCACTAAACGCGATGATATTCGA--CTACA--CGGGAT	1135
Db	563982	TCCGAATGTTATTTGGATTTTGTGATTTACCAACGCGATGATCAATTGATGCTTAAATTAGGCAT	563923
Qy	1136	TTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGCCATTTGGCAAGCGCAGGAA	1195
Db	563922	TATCAAGCCGATATTTGGTATTTCTGATGGGGTATTTGGGTATTTGGCAACGAGGTAA	563863
Qy	1196	CAAGCATGCAAGATGCGTAAGCCCTCATATGCTGTGGGTGTGGGCACAGAAAGCACT	1255
Db	563862	CCCTGACACCATGATAGCTCACACCAATATGATTTATCGGTGCAAGCAACGAAATTTCA	563803
Qy	1256	AGCAGGGGAAGGTATGATTTATTCGCTGGGGGAATTCAGATTCACACACCCACTTCTCTTC	1315
Db	563802	TAAACGCTGCATTTAAATTTGCAACCGCTGGTATCGATACCCACATTTCACTTTATTG	563743
Qy	1316	TCCACACAAATTCCTACCGCTCTAGCCAAATGGCGTTACACCAATGTTTGGAGGCGGCAC	1375
Db	563742	TCCCAACAAAGCAACAACATGCAATTTGAAGTGGCGTTACCAAGTTAAATTTGGTGGGAAC	563683
Qy	1376	AGGTCTCTAGATGGCAGAAATCGCACTACTATCACTCGGGCAATGGAACTTGCAACG	1435
Db	563682	TGGCCCTGCTGATGGTACACACGCAACCACTTTGTACCTTGGCGCATGGTATATGGAAG	563623
Qy	1436	CATGTCGGCGACGACGAAGAGTATTTATGAATGTGGCTTTTGGCAAGGCAATAG	1495
Db	563622	TATGTTTCAAGCGGAGAAAGCTTTGCCGGTAAACGTCGGATTTTGTGTAAGGCAACTG	563563
Qy	1496	CTCTAGCAAAAAACAACTTTAGAACATAGTAGAGCGGGCGGATTTGTTTAAATTCGA	1555
Db	563562	TTCAACCTTAGATCTCTTCGCGTGAGCAAAATTTGAAGCGGGTGCAATTAGTTTAAAAATCCA	563503
Qy	1556	TGAAGACTGGGGCAACACCAAGTGGCATCGATCACTGCTTTGAGCGTGGCAGATGAATA	1615
Db	563502	CGAAGACTGGGTGCAACGCCCTGGCGTATTTGATCTCCCTTTAAAGTAGCAGATGAAT	563443
Qy	1616	CGATGTGCAAGTTTGTATFCCAACCGATACAGTCAATGAGCGAGGTATGTAGATGACAC	1675
Db	563442	GGATATTTCAAGTGGCCATTTACACAGACACGCTAAATGAAAGTGGCTTTTGGGAAGACAC	563383
Qy	1676	CTTAATGCAATGAACGGGCGCGCATCCATGCCATACACATTTGAGGAGCGGTGGAGG	1735
Db	563382	GATGAAGCGATTTGATGGACAGATCATTTACATCTTCATACGGAGGCGCAGGTGGTG	563323
Qy	1736	ACATCACTGATGTATTATCCATGCGGAGCGAGCTCAATATTCTACCTCCTCCACACAC	1795
Db	563322	TCATGCACTTGACATATTAAAGCAGCGATGTTTCAACGTTATCTGCTTTCAACCA	563263
Qy	1796	CCCACATTTCCCTATACCATTAAATACGGTTTGCAGAACATTTAGACATGCTCATGACATG	1855
Db	563262	CCCGACTCGTCGTTTACCRAAAACACCATTTGATGAACATTTGGATATTTGATGTTTG	563203
Qy	1856	CGACCCTTACAAACGCATCCGGAGGATTTTCAATTTTCTTCAAGCGGTATCCGCC	1915
Db	563202	CCATCACTTAGATAAAGCGGTGCGGAGAGCTAGCTTTTCCCATAGTAGCCGTATCCGCC	563143
Qy	1916	CGGCTCTATCGCGCTGAAGATGTCTCATGATATGGTGTGATCGGATGCAAGCTC	1975
Db	563142	TGAACCATTTGCACGCAAGATATTTTGCATGATATGGCGCTCTTCTCAATTATGAGTTC	563083
Qy	1976	GGATTCGAAGCAATGGGGCGTGAGGCGAAGTATTTCTCGAACTTTGGCAGATCGCGGA	2035
Db	563082	AGACTCTAAGCGATGGACGTTATTTGGCGAAGTCGTTATTTCTGTACATGGCAAACTGCAGA	563023
Qy	2036	TAGATATAAAAAAGATTTGGTATAGCTTCTTGAAGATGGCAAGATTAACGATAATTTCCG	2095
Db	563022	TAAGATGAATAATGCAACGTGGTGAGCT-----AGGTAATGAAGGAACGATAACTTCCG	562969
Qy	2096	CATTTAGCGCTACATCTTCCAAATACATCAACCCCGCTTTGTACCCACCGCGTAGCGGA	2155
Db	562968	TATTTAAACGATATATCCGGAATATACCAATCAACCCAGCAATTGCAATGTTATTTGGGA	562909
Qy	2156	GTATATCGGCTCTGTGGAAGAGGCAAGATCGCCGCACTTTGGTGTGTGGAATCTCTGCTT	2215

Db	Seq ID	Score	DB 3	Length	1710	Query Match	22.5%	Score	648.2	DB 3	Length	1710	Best Local Similarity	62.2%	Pred. No.	4.4e-151	Mismatches	633	Indels	6	Gaps	2																						
Db	562908	GCATATTGGCTCGTTAGAAATGGGTAAATCGCATATATCGTTTATGGAAACCGATGTT	562849	2216	TTTGGCGTAAACCCAAAATCGTGATCAAAAGCGGTATGGTGTCTTCTCTGAAATGGG	2275	562848	CTTTGGCGTAAACCTGAATGCGTGAATTAATAAGGCTTTATAGCTATCGAATAATGGG	562789	2276	CGATTCTAAACCGCTGTGTGCCACTCCCCAACCGGTTTATTACCGCGAAATGTTTGGGCA	2335	562788	CGATCCAAATGCCTCAATTCACACACCGCAACCTGTATTTCTACCGTCCAAATGTACGGTGC	2395	2336	TCACGGCAAGCGGAAATTTGACACAGCATCACATTTTGTTCGAAAGTGCCTATGAAAA	562728	ACAAGGCTTAGCAACCGCACAAACAGCAGTAGTATCTTTGTTTCAACGCGCTGAAAAAGC	562669	2396	TGGCGTGAAAGAAAAGCTTGGGCTTTAGAGCGGCAAGTCTTACCGGTCAAAAACCTGCCGTAA	2455	562668	TGATATTTCGTCGAAGTTTCGGTTTACACAAAGAACCAATGCTGTGAAGGCTGCCGAA	562609	2456	CATCACCAAGAAAGACTTCAAGTTCAACGACAAAAACGGCAAAAAATCACCGTCGATCCGAA	2515	562608	CGTAGGTAAAAAGATCTGGTTTCAATGATGTAAACCAACCAATTTACTGTTGATGCTGA	562549	2516	AACCTTCGAGTCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCACCTCGCAAGTGCC	2575	562548	ACGTTATGAAGTTTCGAGTGGACGGAGAGTTAATTACCTGTGCAACCAAGTGATAGGTACC	562489	2576	TCTAGCCAGCGCTACACTTCTTCTTA	2602	562488	ATTGGGTACGGATATTTCTTA	562462
RESULT 12																																												
US-09-543-681A-1857																																												
; Sequence 1857, Application US/09543681A																																												
; Patent No. 6605709																																												
; GENERAL INFORMATION:																																												
; APPLICANT: GARY BRETON																																												
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS																																												
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS																																												
; FILE REFERENCE: 2709.1002-001																																												
; CURRENT APPLICATION NUMBER: US/09/543.681A																																												
; CURRENT FILING DATE: 2000-04-05																																												
; PRIOR APPLICATION NUMBER: US 60/128,706																																												
; PRIOR FILING DATE: 1999-04-09																																												
; NUMBER OF SEQ ID NOS: 8344																																												
; SEQ ID NO 1857																																												
; LENGTH: 1710																																												
; TYPE: DNA																																												
; ORGANISM: Proteus mirabilis																																												
US-09-543-681A-1857																																												
Query Match 22.5%; Score 648.2; DB 3; Length 1710;																																												
Best Local Similarity 62.2%; Pred. No. 4.4e-151;																																												
Matches 1052; Conservative 0; Mismatches 633; Indels 6; Gaps 2;																																												
Qy	912	CAAGAATATCTAAATACCTACGACGCCACCAAGCGGATAAAGTGCCTTAGGAGATACC	971	25	CAAGCTTATCGGATGTTTGGCCCAACACAGCGGATCGTTTCCGATTAGCATATACC	84	972	GATCTTTGGGCGAGAAGTAGAACATGACTATACCACTATGCGGAAGAACTTTAAATTTGGC	1031	85	GAGCTGTTTCTTGAATTTGAAAAGATTTTCCACCACTTATGCGGAAGAGGTCAAATTTGGT	144	1032	GCGGGTAAACTATTCGTTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCTA	1091	145	GGTGGTAAAGTTATTCGTGATGGTATGGGCGCAACGCAAGTTGTTAGTGTGAGTGTGTC	204	1092	GATTTAGTCTACTACTAACCGGATGATTATCGACTACACCGGATTTTACAAAGCCGACATT	1151	205	GATGTTCTGATACCAATGTCATTTATTTAGATTTTGGGCGCATTTGTAAGACGATATT	264	1152	GGGATTAACCAACCGCAAAATCCATGTCATTTGGCAAGCAGGAAACAAAGCATGCAAGT	1211																	

```

RESULT 12
US-09-543-681A-1857
; Sequence 1857, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1857
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-1857

```

Query Match	22.5%	Score 648.2;	DB 3;	Length 1710;
Best Local Similarity	62.2%	Pred. NO. 4.4e-151;		
Matches 1052; Conservative	0;	Mismatches 633;	Indels 6;	Gaps 2;

Qy	912	CAAGAAATATGTAATACTCTACGAGACCACCAAAAGCGGATAAAGTGCCTTTAGGAGATACC	971
Db	25	CAAGCTTATGCGGATATGTTGGCCCAACAACAGGCGATCGTTTGCATTAGCAGATACC	84
Qy	972	GATCTTTGGCAGAGATGAAACATGACTATACCACTATGCGGAAGAACTTAAATTTGGC	1031
Db	85	GAGCTGTTCTTGAAATTGAAAAGATTTCCACCTATGCGGAAGAGTCAAAATTTGGT	144
Qy	1032	GCGGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAACACCCTA	1091
Db	145	GGTGGTAAAGTTATTCTGTATGCTATGGGCCAAAGCCAAAGTTGTAGTCTGAGTGTGC	204
Qy	1092	GATTTAGTTCATCACTAAACGGGATGATTTCGACTACACCGGGATTTACAAAGCCGCATT	1151
Db	205	GATGTTCTGATCACCAAATGCCATTATTTTAGATTTATGGGGCCATTGTAAAGCGAGATATT	264
Qy	1152	GGGATTTAAAAACGGCAAAATTCATGGCATTTGGCAAGCGACGAAACAAGGATCGCAAGAT	1211

Db 265 GGCATTAAAGATGCGCGTATTTGTCGGTATTTGGCAAGGACAGTAATCCAGATGTTCA----- 320
Qy 1212 GGCCTAAGCCCTCATATGCTGCTGGGTGTTGGGCAAGAGCACTAGCAGGAGGAGGTATG 1271
Db 321 -GCCCAA-TGTGGATATTTGTCATTTGGCCCCGGAACAGAGTTGTTGGCGGAGAGGTAAA 378
Qy 1272 ATTATTAACCGTGGGGAATGATTTACACACCACTTCCTTTCTCCCAACAATTCCT 1331
Db 379 ATTGTCATGCTGGTGGTATTTGATACCATATCCACTTTATTTGTCGCAACAAGCCCAA 438
Qy 1332 ACCGCTTAGCCATGGGTATACACCACTTTGGAGCGGCAAGTCTCTGTAGATGCG 1391
Db 439 GAAGGTCTCGTTTCTGGCGTAACCACTTTATTTGGTGGAGGAAAGCGCCCTGTGGCGGT 498
Qy 1392 ACGAATGCGACTATCATCTCCGCGGCAAAATGGAATTTGCACCGCATGTTGGCGCAGCA 1451
Db 499 ACTAATGCAACACAGTTACCCCGGTATTTGGAAATATGTACCGCATGTTAGAGCGGTG 558
Qy 1452 GAAGAGTATTTCTAATGATGTTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACA 1511
Db 559 GATGAAATACCTAATTAATGTTGGGTTATTTGGTAAAGTTGTGTGTCAGTCAGCCGCAAGCA 618
Qy 1512 CTTGTAGAACAGTAGAGCGGCGCGATTTGTTTAAATTTGATGAAGACTGGGCGACA 1571
Db 619 ATCCGCGAACAATAACAGCGCGGTCTATAGGTCTTAAATAACATGAAGACTGGGCGCA 678
Qy 1572 ACACCAATGCGATCGATCTACTGCTTGTAGCGTGGCAGATGAATACGATGTCGAAGTTGT 1631
Db 679 ACGCCAATGGCAATTCACATTTGCTTAAATGCTGCGCGATGAATGATGTACAGTGGCT 738
Qy 1632 ATCCACACCGCATACAGTCAATGAGGACAGTTATGTAGTAGCACCTTAAATGCAATGAAC 1691
Db 739 ATTCACTCTGACACTTAAATGAAGTGTGTTTATGAAGAGACAGTAAAAAGCCATTGCC 798
Qy 1692 GGGCGCGCATTCATGCCCTACCAATTGAGGAGGGGTGAGGACACTCACCTGATGTT 1751
Db 799 GGTGAGTGTATCCATGATTTCCATACCAAGGCGGAGTGGTGTGTCAGCCCTGATGTT 858
Qy 1752 ATCACCATTGGCAGCGAGCTCAATATTTCTCCCTCCACACCCCACTATTCCCTAT 1811
Db 859 ATCAAGTCGGTAGAGAGCCCAATATTTTCTGTCATCAACCAACCAAGATGCTTAT 918
Qy 1812 ACCAATTAATCGTTGCAGAACACTTAGACATGCTCATGATGTCACCACTAGACAAA 1871
Db 919 ACCAATTAATCGTTGCAGAGCACTTGTATGTTGATGTTCTGTCTCATCACTCGATCCC 978
Qy 1872 CGCATCCGCGAGGATTAACAATTTCTCAAGCCGTATCCGCCCGGCTCTATCGCGGT 1931
Db 979 TCTATTCTGAAGATGTGGCATTTGCTGAATCTGTTATTCGTCGCAAAACCAATGCTGCA 1038
Qy 1932 GAAGATGCTCTCATGATATGGTGTGATCGGATGACAACTCGGATTCGCAAGCAATG 1991
Db 1039 GAAGATATCTTACATGATATGGGGCAATTTCCGTGATGTCGTGATGTCAGACTCAAGCCATG 1098
Qy 1992 GGGCGTCAGCGCAAGTATTTCTCGAACTTTGGCAGACTCGGCTGAAGATAAATAAAGAA 2051
Db 1099 GGAAGTTCGAGAGATTTATCTTACGCACTTTGGCAGTGTGCATATAATGAATGCAA 1158
Qy 2052 TTTGCTAAGCTTCTGAAGATGGCAAGATAACGATAATTTCCGATTTAAGCGCTACATC 2111
Db 1159 CGAGGCACTTAGCGGTGATAGCGCAGATAATGATAATAATCGTATTAAACGTTATATC 1218
Qy 2112 TCCAATACACTATCAACCCGCTTTGACCAACCGGTGAGGAGTATATCGGCTCTGTG 2171
Db 1219 GCTAATAACAGATTAATTCGCGCACTGGCGACATGGCATTTGCTATACCGTGGGATCAATA 1278
Qy 2172 GAAGAGCAAGATCGCGACTTGTGTGTGGAATCTGCTTTTGGCGTAAACCC 2231
Db 1279 GAAAAGGTAACCTTGCGGATATCGTGCTATGGGATCTGCTTTCTTTGGCGTCAACCG 1338
Qy 2232 AAAATCGTGATCAAGGCGGTATGGTGTCTTCTGAAATGGGCGATTTCAACCGCTCT 2291

Db 1339 GCATTATCATAAAGGTGGTATGTCGCTATGCGCCAAATGGGGGATATTTATGCGGCT 1398
Qy 2292 GTGCCACTCCCAACCGGTTTATTTACCGCGAAATGTTTGGGCATCACGCAAGGCGAAA 2351
Db 1399 ATTCAACACCGCAACCGGTTCAITTTATCGTCCAATGATATGCTGTTTAGGAAAGCCAAA 1458
Qy 2352 TTGACACAGCATCACTTTTGTGTTTCCAAAGTCGCTATGAAAAATGGCGTGAAGAAAG 2411
Db 1459 TATCAACGTCGATGATCTTTATGTCAAAAGCGGGTATTTGAGCGGGAGTGCAGAAAAA 1518
Qy 2412 CTGGCTTTAGAGCGCAAGTTCTTACCGGTCAAAATCTGCGTAACATCAACCAAGAAAGAC 2471
Db 1519 TTAGGCTTTAAAGCTTAATTTGTCGTGTGGAGGCTGTGTCATATCAAAAGCTTCG 1578
Qy 2472 TTCAAGTTCAACGACAAAACCGCAAAAAATCACCGTCGATCCGAAAAACCTTTCCGAGGCTTT 2531
Db 1579 ATGATCCACAATAACTATGTTCTCATATCGAATTAGATCCACAACCTTACATTGTAAA 1638
Qy 2532 GTAGATGCAAACTCTGCACCTCTAAACCCACCTGCGAAGTGCCTTAGCCCAAGCGCTAC 2591
Db 1639 GCGGATGTTGTACCACTGTTTGTGAGCCAGCGACTGAATTACCGATGGCTCAACGCTAT 1698
Qy 2592 ACTTTCTTCTA 2602
Db 1699 TTCTTATTTTA 1709

RESULT 13
US-08-967-513-1
; Sequence 1, Application US/08967513
; Patent No. 5783436
; GENERAL INFORMATION:
; APPLICANT: Robert P. Hausinger
; TITLE OF INVENTION: Mutant Urease and Method
; TITLE OF INVENTION: of Use For Determination
; TITLE OF INVENTION: of Urea
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,513
; FILING DATE: 11-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/687,645
; FILING DATE: July 26, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-309
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: No. 57834366
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2400
; TYPE: nucleotides
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: cdna

HYPOTHETICAL: No	ANTI-SENSE: No	ORIGINAL SOURCE:	ORGANISM: Klebsiella aerogenes	STRAIN: CG253	INDIVIDUAL ISOLATE:	CELL TYPE: N/A	FEATURE:	NAME/KEY: cdna encoding mutant urease	NAME/KEY: 'H219Q	LOCATION: Modification at position 1312 to	LOCATION: glutamine	IDENTIFICATION METHOD: Sequencing	OTHER INFORMATION:	Query Match	21.4%;	Score 617.2;	DB 2;	Length 2400;	Best Local Similarity 60.3%;	Pred. No. 2.7e-143;	Mismatches 0;	Mismatches 683;	Indels	Gaps	1;																																																							
883	GTAGGAGACTCCCATGAAATGAAAAAACAAGAATATGTAATACCTACGGACCCACCA	942	645	GGAGTAAACGATGAGTAATATTTCACGCCAGGCCTATGCCGATATGTCGGCCCAACG	704	943	AAGCGATAAAGTCGGCTTAGGAGATACCGATCTTTGGCGAAGAAGTAGAACAATGACTATA	1002	705	TCGGCGCAAGGTGCGCTGCGCATACCGAGCTGTGGATCGAGGTGGAGACGATTTGA	764	1003	CCACCTATGCGGAAGACTTAAATTTGGCGGGTAAAACTATCCGTGAGGTTATGGTC	1062	765	CCACCTACGGGGAAGAGGTCAAAATTCGGCGCGGCAAAAGTATCCGACGGCATGGGCC	824	1063	AGAGCAATAGCCCTGATGAAAAACACCCCTAGATTAGTCATCACTAACCGCCATGATTATCG	1122	825	AGGACAGATGCTGGCGCGCAGCATGTGTGCACCTGGTGCTCACCAACGCTTGATCGTCG	884	1123	ACTACACCGGATTTTAAAGACCGACATTTGGGATTTAAAAAACGGCAAAATCCATGGCAATG	1182	885	ATCACTGGGGGATCGTTTAAAGCCGATATCGGCGTGAAGACGGCGGATCTTCGCCATCG	944	1183	GCAAGGCAGAAACAAGGACATCGAAGATGGCGTAAGCCCTCATATGGTCTGGTGGTGCG	1242	945	GCAAGCGCGCAACCCGACATCCAGCCCAACGTCACC-----ATCCCCATCGGCGCTG	998	1243	GCAAGAAGCACTAGCAGGGGAAGGTATGATTATACCGCTGGGGGAATCGATTCACACA	1302	999	CGACGAGATGATCGCGCGAAGGAAAATTTGTACCGCGCGGAGATCGATACCCATA	1058	1303	CCCACTTCCTTTCTCCACAACAATTCCTACCGCTTAGCCAATGGCGTTTACAACCATGT	1362	1059	TTCACTGGATCTGTCCGACGAGCGGAGAGCGCTGGTCTCTGGCGTGACCAACATGG	1118	1363	TTGAGCGGCAAGGTCTGTAGATGGCAGAAATGGGACTACTATCACTCCGGGCAAT	1422	1119	TCGGCGCGGACACCGCCCGCGCGCGGCCCAATGTCACCATCTGCACCCCGGGGCCGT	1178	1423	GGAACCTTGCAACCGATTTGCGCGCAGCAGAAGATTTCTATGAATGTGGGCTTTTGG	1482	1179	GGTATATCTCACGCATGTCTGACGCGCGCGACAGCTCCCGCTCAATATCGGCTCTGCTG	1238	1483	GCAAGGCAATAGTCTTAGCAAAAAACAATTGTAGAACAAAGTAGAAGCGGGCGCGATTTG	1542	1239	GCAAGGAAAACGTTTCTCAGCGGATGCCCTGCGCGAGCAGAGGTGGCGGAGGCGTTATTG	1298	1543	GTTTTAAATTGCATGAAGACTGGGGCAACAACCAAGTGGGATCGATCACTGCTTGAGCG	1602	1299	GCCTGAAGATCCAAGAGGACTGGGGGCGCACCCCGGGCGGATCGACTGTGGCGTTAACCG	1358	1603	TGCGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTT	1662	1359	TCGCCATGAATGGACATCCAGGTGCCCTGCGACAGCGACACCTCGAATGAATCCGGTT	1418	1663	ATGTAGATGACACCCCTAAATGCAATGAAGCGGCGCGCCATCCATGCTACCAATGAGG	1722

RESULT 14
US-08-687-645B-1
; Sequence 1, Application US/08687645B
; Patent No. 5846752
; GENERAL INFORMATION:
; APPLICANT: Robert P. Hausinger
; TITLE OF INVENTION: Mutant Urease and Method
; TITLE OF INVENTION: of Use For Determination
; TITLE OF INVENTION: of Urea

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: wordperfect 5.1
CURRENT APPLICATION DATA:
FILING DATE: July 26, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 5846752e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2400
TYPE: nucleotides
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE:
DESCRIPTION: cdna
HYPOTHETICAL: No
ANTI-SENSE: No
ORIGINAL SOURCE:
ORGANISM: Klebsiella aerogenes
STRAIN: CG253
INDIVIDUAL ISOLATE:
CELL TYPE: N/A
FEATURE:
NAME/KEY: cdna encoding mutant urease
NAME/KEY: 'H219Q
LOCATION: Modification at position 1312 to
LOCATION: glutamine
IDENTIFICATION METHOD: Sequencing
OTHER INFORMATION:
US-08-687-645B-1
Query Match 21.4%; Score 617.2; DB 2; Length 2400;
Best Local Similarity 60.3%; Pred. No. 2.7e-143;
Matches 1048; Conservative 0; Mismatches 683; Indels 6; Gaps 1;
Qy 883 GTAAGGACCTCCCATGAAATGAAAAACAGAAATATGTAATACCTACGACCCACCA 942
Db 645 GGAGGTAACAGTACGATGATTAATTTCCAGCCAGGCTATGCCGATATGTTGGCCCCACCG 704
Qy 943 AAGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAACAATGACTATA 1002
Db 705 TCGGGACAAAGTGGCTTGCAGATACCGAGCTGTGGATCGAGTGGAGACGATTTGA 764
Qy 1003 CCACCTATGGCGAAGAACTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGTTC 1062
Db 765 CCACTACGGGAGAGGTCAAATTTGGCGGGCGCAAGTATCGCGACGGCATGGGCC 824
Qy 1063 AGAGCAATAGCCCTGATGAACACCCCTAGATTAGTTCATCACTAACCGGATGATTTCG 1122
Db 825 AGGACAGATGCTGGCCCGCGACTGTGCGACCTGTGCTCAACCGGTTGATCGTCG 884

Qy 1123 ACTACACGGGATTTTAAAGCCGACATTGGGATTTAAAAACGGCAAAATCCATGGCAATTG 1182
Db 885 ATCACTGGGGATCGTTAAGGCCGATATCGGCGTGAAGACCGCGCGGATCTTCGCGCATCG 944
Qy 1183 GCAAGGCAAGGAAACAAAGGACATGCGAAGATGGGTAAAGCCTCATATGGTCTGGTGGTGG 1242
Db 945 GCAAGCCGGCAACCCCGACATCCAGGCCAAAGTCAAC-----ATCCCATCGGCGCTG 998
Qy 1243 GCACAGAAAGCACTAGCAGGGGAAGTATGATTTATTAACCTGGGGGAATCGATTTCAACA 1302
Db 999 CGACGGAAGTATCGCGCGGAAGGAAAAATTTGTACCGCGCGGGATCGATACCCATA 1058
Qy 1303 CCCACTTCTTCTCCACAACAAATTCCTACCGCTAGCCCAATGCGGTATCAACCATGT 1362
Db 1059 TTCACTGGATCTGTCCGAGCAGGCGGAAGAGCGCTGTCTCTGGCGTGAACCATCGG 1118
Qy 1363 TTGGAGGCGGCACAGCTCTGTAGATGGCAAGTGCAGAACTGCGACTACTATCACTCCGGGCAAA 1422
Db 1119 TCGCGGCGGCACCGCGCGCGCGGCGGCCACCCATGCCACCACTGCAACCCCGGCGCT 1178
Qy 1423 GGAACCTTGCAACCGCATGTTGCGCGCAGCAGAAAGATTAATTTATGAATGTGGGCTTTTGG 1482
Db 1179 GGTATATCTCAGCATGCTGCAGGCGCGCACAGCCTGCGGTCAATATCGGCTCTCTGG 1238
Qy 1483 GCAGAGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAAAGTAGAAGCGGGCGCGATTG 1542
Db 1239 GCAAGGGAAACGTTTTCTCAGCCGGATGCCCTCGCGAGCAGGTGGCGGCGGCTTTATTG 1298
Qy 1543 GTTTTAAATTCATGAAGACTGGGGCACAAACCAAGTCCGATCGATCACTGCTTCAGCG 1602
Db 1299 GCCTGAAGATCAAGAGGACTGGGGCGCACCCCGCGCGGATTCGACTGTGGGTTAAACG 1358
Qy 1603 TGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATAGAGCGAGTT 1662
Db 1359 TCGCGATGAATGACATCCAGGTGCGCTGCACAGCGACACCCCTGAATGAATCGGTT 1418
Qy 1663 ATGTAGATGACACCCCTAAATGAAGACGGGGCGCCCATCGATGCTACCACTTGAAGG 1722
Db 1419 TTGTGGAAGACACCCCTCGCGCCATCGGGCGGCGCACCATCCACACCTTCCATACCGAAG 1478
Qy 1723 GAGCGGTGGAGGACACTCACCTGATGTTATCACCATGCGCGCGAGCTCAATATTCTAC 1782
Db 1479 GGGCGGCGGGCGCCATGCGCGGACATCATCACCCTGCGCCACCCGAAACATTTTGC 1538
Qy 1783 CCTCTCCACACCCCTCACTATTCCCTATACCAATTAATACGGTTGCAGAACCTTAGACA 1842
Db 1539 CGTGTCCACCAACCCCAAGCTGCTTACACCTCAACACCATCGATGAACATCTCGATA 1598
Qy 1843 TGCTCATGACATGCCACCACTTAGACAAACGGATCGCGAGGATTTACAATTTCTCAAA 1902
Db 1599 TGCTGATGTTCTGCCACCATCTGGACCCGACATCGCCGAGGACGTGGGCTTTGCGGAGT 1658
Qy 1903 GCCGTATCGGCCCGCTTATCGCGGCTGAAGATGTCCTCCATGATATGGGTGATCG 1962
Db 1659 CGCGCATTTGCGCGGAAACCATCGCTGCGGAAGACGTGTGACACATCTCGGCGCTTCT 1718
Qy 1963 CGATCACAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGCAAGTGATTTCTCGAATCT 2022
Db 1719 CGCTCACCTCTCCGATTCGAGGCCATGGGCGCGCTCGGGGAAGTGAATTTCTCGCACCT 1778
Qy 2023 GCGAGCTCGGATGAAGATAAAAAAGAAATTTGGTAAAGCTTCTGAAGATGGCAAGATA 2082
Db 1779 GCGAGTGGCGCATCGCATGAAGGTGCGAGCGGAGCGCTGGCGGAGGAGACCGGGGATA 1838
Qy 2083 AGGATAATTTCCGCAATTAAGCCCTACATCTCCAAATACACTATCAACCCCGCTTTGACCC 2142
Db 1839 AGCAACAACCTTCGCGTGAAGCGCTACATCGCAAAATACACCATCAACCCCGCGCTGACCC 1898
Qy 2143 ACGCGGTGAGCGAGTATATCGCTCTGTGGAAGAGGCGCAAGATCGCGCACTTGGTGGTGT 2202
Db 1899 ACGGATCGCACAGAGTCCGATTCATTTAGGTGGTGAAGCTGGCTGACCTCGTGGTCT 1958
Qy 2203 GGAATCTCGCTTTTGTGGCGTAAAAACCCCAAAATCGTGATCAAAAGCGGATGCGTGTCT 2262

Db 1959 GGTACACAGCTTCTTTCGGCGTGAACCGGCCACCGTATCAAGGGCGCATGATCGCCA 2018
Qy 2263 TCTCTGAATGGGCGATTTCTAACGGCTCTGTGCGCACTCCCAACCGGTTTATTACCGCG 2322
Db 2019 TCGCGCGATGGGCGATATCAATGCTCTATTCCGACCCCGCAGCGGTGCACTACCGCC 2078
Qy 2323 AATGTTTGGGTCATCGGCAAGGGAAATTTGACACCGCATCACATTTTCTTCCAAAG 2382
Db 2079 CGATGTTTGGCGCTGGGCGAGCGCCGCCATCACTGCGGCTCACCTTCTGTGCGAGG 2138
Qy 2383 TCGCTATGAATGGCGTGAAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCA 2442
Db 2139 CGGGCGAGCAATGGGCTTGGCGAGCGGCTGAACCTTGGCGAGCGGATCGCGTGGTGA 2198
Qy 2443 AAAATGCGGTAAATCACCAAGAAAGACTTTCAAGTTCAACGACAAAAACGCAAAAAATCA 2502
Db 2199 AAGCTGCGGTACGGTCAGAAAGCCGACATGGTGCACAAACAGTCTGCGAGCTTAACATCA 2258
Qy 2503 CGGTGATCCGAAAACTTTCGAGTCTTTTGTAGATGGCAAACTCTGCACTCTTAAACCCA 2562
Db 2259 CCGTCAAGCCGACACTATGAGGTGCGGGTGGATGGCGAACTTATCACCAAGCGAGCGG 2318
Qy 2563 CTTGCGAAGTGCCTTAGCCAGCGCTACACTTTCTTCTAGGCAATGCCCCCTTT 2619
Db 2319 CAGAGTTCTGCGCATGGCGAAGATATTTCTGTTTAAAGGAGCGGATGCTTT 2375

RESULT 15

US-09-489-039A-2045
; Sequence 2045, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2045
; LENGTH: 1878
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2045

Query Match 21.3%; Score 612.8; DB 3; Length 1878;
Best Local Similarity 60.8%; Pred. No. 3e-142;
Matches 1028; Conservative 0; Mismatches 657; Indels 6; Gaps 1;
Qy 900 AATGAAAGAAACAGATATGTAATACCTACGGACCCACCAAGGCGATAAGTGGCG 959
Db 10 AATATTTCAGCCAGGCTATGCGATATGTTGCGGCCACCGTGGCGCAAGGTGCGC 69
Qy 960 TTAGGAGATACCGATCTTTTGGCGAGAGTAGAATCACTATACCACTATGGCGAAGAA 1019
Db 70 CTGGCGATACCGAGCTGTGGATCGAGGTGAGGACGATCTGACCACTACGGGGAAGAG 129
Qy 1020 CTTTAAATTTGGCGCGGTAAACCTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT 1079
Db 130 GTCAATTTGGCGCGGCAAGTATCCGCGAGCGCATGGCGATGGCGGCGAGATGCTGCCC 189
Qy 1080 GAAAAACCCCTAGATTAGTATCATCACTAACCGGATGATTATGACTACACCGGATTTAC 1139
Db 190 GCGCACTGTGCTCACTCTGTGCTCACCAAGCGTTGATCGTCACTACCTGGGGGATCGTT 249
Qy 1140 AAGCGCATTTGGGATTAAGAAAGCGCAAAATCCATGGCATTTGCGAGGCGAGGAAACAG 1199
Db 250 AAGGCGGATATCGGCTGGAAGGACGCGCGGATTTCTGCCATCGGCAAGCGGTAACCCC 309
Qy 1200 GACATGAAGATGGCGTAAGCCCTCATATGTTGTTGGTGTGGGCGACAGAAGCACTAGCA 1259

Db 310 GATATCCAGCCCAACGTACCC-----ATCCCATCGGTGCTCGACGGAAGTGTGCC 363
Qy 1260 GGGGAAGGTATGATTTATCCGCTGGGGAAATCGATTCAACACCCACTTCTTCTTCCA 1319
Db 364 GCGGAAGGAAATTTGTACCGCGCGGGATCGATACCATATTCACTGGATCTGTCCG 423
Qy 1320 CACAAATTTCCCTACCGCTTAGCCATAGCGGTATCAACCATGTTTGGAGCGGACAGGT 1379
Db 424 CAGCAGGCGGAAGAGCGGCTGTCTCCCGCGTGAACCAACATGGTTCGGCGCGACCGCG 483
Qy 1380 CCGTGTAGATGGCAGAAATCGGACTACTATCACTCCGGGCAAAATGGAATTTGACCGCATG 1439
Db 484 CCGCGCGGGGACCCATGCGCACCTGCAACCCCGGGCCCGTGGTATATCTCAGCGATG 543
Qy 1440 TTGCGCGCAGCAGAAAGAGTATTCTAATGTGGGCTTTTGGGCAAAAGCAATAGCTCT 1499
Db 544 CTGACGCGCGCGACAGCGCTGCGGTCAATATCGGCTGCTGGGCAAGGCAACGTTTCT 603
Qy 1500 AGCAAAAAAACAATTGTGTAGAACAGTAGAGCGGGCGCGATTTGTTTAAATTTGCATGA 1559
Db 604 CAGCGGATGCGCTGCGGAGCAGGTGGCGGGCGGCTTATTGGSCCTGAAGATCCATGAG 663
Qy 1560 GACTGGGCGCAACACCAAGTGCAGTCACTGCTTCAGCGGTGGCAGATGAATAGAT 1619
Db 664 GACTGGGCGCGCACCCCGCGCGCATCGACTGTGCGTTAACCGTCCCGCATGAATGAGAC 723
Qy 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGGTTTATGTAGATGACACCTA 1679
Db 724 ATCCAGTTCGCTGTCACAGCGACACCTGATGATCGGTTTGTGGAGACACCTC 783
Qy 1680 AATCAATGAACGGCGCGCATCATGCTTACCACTTACCAATTTAGAGGAGCGGTGAGGACAC 1739
Db 784 GCGGCAATCGGCGGCGACCATCATCACACCTTCCATACCGAAGGGCGCGCGCGCAT 843
Qy 1740 TCACCTGATGTTATCACCATGGCAGCGAGTCAATATTCTACCTCTCCACCAACCCC 1799
Db 844 GCGCGGACATCATCACCGCTCGCGCCACCCGACATTTTGGCGTCTCCACCAACCCA 903
Qy 1800 ACTATTCCTTATACCAATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCAC 1859
Db 904 AGGTGCGCTTACACCTCAACACCATCGATGAATCTCGATATGCTGATGTTGCTGCCAC 963
Qy 1860 CACCTAGACAAAACGATCGCGAGGATTTACAATTTTCTCAAAGCGGTATCCGCCCGCGC 1919
Db 964 CATCTGGACCCGACATCGCGAGACGTGGCTTTTGGCGAGTCCGCGCATTCGCCGGAA 1023
Qy 1920 TCTATCGCGGTGAAGATGCTCTCATGATATGGTGTGATCGCGATGACAGCTCGGAT 1979
Db 1024 ACCATCGCTGGGAAGACGTGTGACGATCTCGGTGCTTCTCGCTCACCTCCTCCGAT 1083
Qy 1980 TCGCAAGCAATGGGCGTGCAGGCGAAGTATTCCTCGAACTTTGGCAGATGCGGATAAG 2039
Db 1084 TCGCAGGCGATGGGCGGCTCGGGGAAGTATTTCTCGCACCTGGCAGGTGGCGCATCGC 1143
Qy 2040 AATAAAAAGAAATTTGGTAAGCTTCTCGAAGATGGCAAGATAACGATAATTTCCGCAAT 2099
Db 1144 ATGAAGTGCAGCGGAGCGCTGGCGAGGAGACCGGGGATAACGACAACTTCCGCGTG 1203
Qy 2100 AAGCGCTATCTTCCAAATACATTAACCCCGCTTTTGAACCAACCGCTGAGCGAGTAT 2159
Db 1204 AAGCGTACATCGCCAAATACCACTCAACCGCGGCTGACCAACCGCATCGCACAGAA 1263
Qy 2160 ATCGCTCTGTGAAGAGGCAAGATCGCGCATTTGGTGTGTTGTAATCTCCTGAAATGGCGCAT 2219
Db 1264 GTCGGATCCATTGAGGTGGGTAAGCTGTGACTCGTGGTCTGGTCAACAGCGCTTCTTC 1323
Qy 2220 GCGGTAAAAACCAAAATCGTGTATCAAGCGCGTATGGTGTCTTCTGAAATGGCGCAT 2279
Db 1324 GCGGTGAACCGGCGCACCGTGTATCAAGCGGCGCATGATCGCCATCGCGCGGATGGCGCAT 1383
Qy 2280 TCTAACCGCTGTGTGCCCACTCCCAACCGGTTTATTATCCGCGAAATGTTTGGGCAATCAC 2339

```

Db      1384 ATCAATGCCTCTATTCCGACCCCGCAGCGCGTGCACTACCGCCCGATGTTTGGCGCGTG 1443
Qy      2340 GGCAAGCGGAAATTTGACACCGAGCATCACTTTTTCCTCAAAAGTCGCTATGAAAAATGGC 2399
      ||||| ||
Db      1444 GGCAGCGCCCGCCCATCACTGCGGCTCACCTTCCTGTGCGAGCGCGGCGCAATGGC 1503
      ||||| ||
Qy      2400 GTGAAAGAAAAGCTGGGCTTAGAGCGGCAAGTTCTACCGGTCAAAAAGTGGCGTAACATC 2459
      ||||| ||
Db      1504 GTTCCCGAGCGGCTGAACCTGCGCAGCGCGATCGCCGTGTGAAAGGCTGCCGTACGGTG 1563
      ||||| ||
Qy      2460 ACCAAGAAAGACTTCAAGTTCAACGACAAAAACGGCAAAAATCACCGTCGATCCGAAACCC 2519
      ||||| ||
Db      1564 CAGAAAGCCGACATGGTGCAACAGCCTGCGAGCTAACATCACCGTCGACGCCGAGACC 1623
      ||||| ||
Qy      2520 TTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAAGTGCCTCTA 2579
      ||||| ||
Db      1624 TATCAGGTGCGGTGGATGGCGAACTTATCACAGCGAGCGCGCAGAGGTTCTGCCGATG 1683
      ||||| ||
Qy      2580 GCCCAGCGCTA 2590
      ||||| ||
Db      1684 GCGCAACGATA 1694
      ||||| ||

```

Search completed: August 10, 2006, 02:19:22
 Job time : 542 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:39:48 ; Search time 4662 Seconds

(without alignments)
7598.724 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2880.6

Sequence: 1 rrgagattttccarctt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY_NUC

Gapop 3.0 , Gapext 3.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main.*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2880.6	100.0	2883	3	US-09-904-994B-1
2	2228	77.3	2452	3	US-09-904-994B-13
3	2158.6	74.9	2405	3	US-09-904-994B-4
4	2158.6	74.9	2407	3	US-09-904-994B-10
5	1948	67.6	2183	3	US-09-904-994B-7
6	1134.4	39.4	8407	8	US-10-639-273-1
7	940.8	32.7	1710	8	US-10-639-273-4
8	902.4	31.3	1719	9	US-10-476-313-10
9	886.4	30.8	1710	8	US-10-335-977-3849
10	885.8	30.8	1815	3	US-09-895-913A-251
11	884.8	30.7	1710	8	US-10-282-122A-22427
12	880	30.5	1710	10	US-10-500-447A-5
13	876.8	30.4	1717	9	US-10-476-313-11
14	860.6	29.9	2385	3	US-09-402-100-1
15	665.4	23.1	1830121	8	US-10-329-670-1
16	665.4	23.1	1830121	9	US-10-158-865-1
17	665.4	23.1	1830121	10	US-10-981-687-1

18	653.6	22.7	1698	8	US-10-282-122A-33102	Sequence 33102, A
19	650.4	22.6	2341	8	US-10-282-122A-32950	Sequence 32950, A
20	636	22.1	2439	8	US-10-282-122A-34390	Sequence 34390, A
21	624	21.7	3919	10	US-10-795-159-460	Sequence 460, App
22	624	21.7	85814	10	US-10-795-159-460	Sequence 23823, A
23	612.2	21.3	1701	8	US-10-282-122A-23823	Sequence 22026, A
24	612.2	21.3	1719	8	US-10-282-122A-22026	Sequence 22026, A
25	604.6	21.0	1704	8	US-10-282-122A-13582	Sequence 13582, A
26	599	20.8	1719	8	US-10-282-122A-41843	Sequence 41843, A
27	598.4	20.8	1704	8	US-10-282-122A-11454	Sequence 11454, A
c	593.6	20.6	5966	2	US-08-781-986A-22	Sequence 22, Appl
c	593.6	20.6	5966	2	US-10-329-624-22	Sequence 22, Appl
30	586.4	20.4	1698	8	US-10-282-122A-14054	Sequence 14054, A
31	585.6	20.3	1704	8	US-10-282-122A-14054	Sequence 14054, A
32	582	20.2	1701	8	US-10-282-122A-32068	Sequence 32068, A
33	572.4	19.9	8729	6	US-10-282-122A-30559	Sequence 30559, A
34	569.8	19.8	8729	6	US-10-114-170-258	Sequence 57, Appl
35	569.8	19.8	87563	6	US-10-114-170-57	Sequence 20639, A
36	568.8	19.7	1707	8	US-10-282-122A-20639	Sequence 20660, A
37	568.8	19.7	1707	8	US-10-282-122A-20660	Sequence 3828, Ap
38	563.8	19.6	3164	10	US-10-793-626-3828	Sequence 3823, Ap
39	563.8	19.6	3234	10	US-10-793-626-3853	Sequence 117, App
40	562.8	19.5	1716	10	US-10-793-626-117	Sequence 2754, Ap
41	562.8	19.5	1722	8	US-10-724-972A-2754	Sequence 14888, A
42	562.2	19.5	1713	8	US-10-282-122A-14888	Sequence 1, Appl
43	545.2	18.9	3309400	3	US-09-738-626-1	Sequence 99, Appl
44	543.2	18.9	1710	3	US-09-738-626-99	Sequence 8259, Ap
45	542.6	18.8	1716	8	US-10-282-122A-8259	

ALIGNMENTS

RESULT 1

US-09-904-994B-1
; Sequence 1, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1e1 N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodsequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2883
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (206)..(886)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (897)..(2603)
US-09-904-994B-1

Query Match	100.0%	Score	2880.6	DB	3	Length	2883
Best Local Similarity	100.0%	Pred.No.	0				
Matches	2883	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	RGGAGATTTTCCARCACTTCAAGCACATATTGATCTCTGTGTGGTGGTAAATTCR	60				
Db	1	RGGAGATTTTCCARCACTTCAAGCACATATTGATCTCTGTGTGGTGGTAAATTCR	60				
Qy	61	ACTTGTAAATCTATTATTAATTTTTTAAATTAATCTTATTATCATATATAATATTA	120				
Db	61	ACTTGTAAATCTATTATTAATTTTTTAAATTAATCTTATTATCATATATAATATTA	120				
Qy	121	TTACTTATATAAAGTTAATAAAGTACGAAATTAGGACTATAATCCCATTCCTT	180				
Db	121	TTACTTATATAAAGTTAATAAAGTACGAAATTAGGACTATAATCCCATTCCTT	180				

181 TAAATTTAAACAAGGAGTAATAGGTGAATCTCACACCCAAAGACGAAAGAAAGTTCTT 240
181 TAAATTTTAAACAAGGAGTAATAGGTGAATCTCACACCCAAAGACGAAAGAAAGTTCTT 240
241 GTTATATATTGCGGGCGAAGTGGCTAGAAAGCGCAAGACGAGGGCTTAAAGCTCAACCA 300
241 GTTATATATTGCGGGCGAAGTGGCTAGAAAGCGCAAGACGAGGGCTTAAAGCTCAACCA 300
301 ACCGAAGCCATTGCTTTACATTAGTGGCCCATCGAACCAGATGAGCACTTCAAAGCGGCGAAGTGAA 540
301 ACCGAAGCCATTGCTTTACATTAGTGGCCCATCGAACCAGATGAGCACTTCAAAGCGGCGAAGTGAA 540
361 AACCGTTGCCAGCTTATGGAAGAGTGATGCACTTTTGAAGAAAGATGAAGTAATGCC 420
361 AACCGTTGCCAGCTTATGGAAGAGTGATGCACTTTTGAAGAAAGATGAAGTAATGCC 420
421 CGGGTGGGTAAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCCTGATGTCAGAA 480
421 CGGGTGGGTAAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCCTGATGTCAGAA 480
481 ACTTGTAACTGTGAAATTTGGCCCATCGAACCAGATGAGCACTTCAAAGCGGCGAAGTGAA 540
481 ACTTGTAACTGTGAAATTTGGCCCATCGAACCAGATGAGCACTTCAAAGCGGCGAAGTGAA 540
541 ATTTGGTTGGCATAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAACCGAACTTGAGGT 600
541 ATTTGGTTGGCATAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAACCGAACTTGAGGT 600
601 TACTAAATGAAGGGCTAAATCTTGCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAA 660
601 TACTAAATGAAGGGCTAAATCTTGCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAA 660
661 CAAGGCACTAAATTCGATCGTGAAGAAAGCCCTATGGCAAGCCCTAGATATTCCTCTGTG 720
661 CAAGGCACTAAATTCGATCGTGAAGAAAGCCCTATGGCAAGCCCTAGATATTCCTCTGTG 720
721 CAACACGCTACGCAATTTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTGTG 780
721 CAACACGCTACGCAATTTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTGTG 780
781 CAGTAAAGAAAGTGGCATGAACGGGCTTGTGAATTAACATCGCGGATGAACGCCATAA 840
781 CAGTAAAGAAAGTGGCATGAACGGGCTTGTGAATTAACATCGCGGATGAACGCCATAA 840
841 ACATAAAGCGCTTCACAAGGCGAAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGA 900
841 ACATAAAGCGCTTCACAAGGCGAAATCTCAGGATTTATCAAGTAAGGAGACTCCCATGA 900
901 AAATGAAAAAACAAGAAATATGTAATACCTACGACCCCAACCAAGGGCGATAAAGTGCGCT 960
901 AAATGAAAAAACAAGAAATATGTAATACCTACGACCCCAACCAAGGGCGATAAAGTGCGCT 960
961 TAGGAGATACCGATCTTTGGGCAGAGTAGAACAATGACTATACACCTATGGCGAAGAAC 1020
961 TAGGAGATACCGATCTTTGGGCAGAGTAGAACAATGACTATACACCTATGGCGAAGAAC 1020
1021 TTAATTTGGCGGGTAAACTACTCCGTAGGGTATGGGTGAGCAATAGCCCTGATG 1080
1021 TTAATTTGGCGGGTAAACTACTCCGTAGGGTATGGGTGAGCAATAGCCCTGATG 1080
1081 AAAACACCTTAGATTTAGTCACTACACGCGATGATTTATGACTACACCGGGATTTACA 1140
1081 AAAACACCTTAGATTTAGTCACTACACGCGATGATTTATGACTACACCGGGATTTACA 1140
1141 AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCAATTTGGCAAGGAGGAAACAAG 1200
1141 AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCAATTTGGCAAGGAGGAAACAAG 1200
1201 ACATGCAAGATGGCGTAGCCCTCATATGGTCTGGGTGGGACAGCAAGCACTAGCAG 1260
1201 ACATGCAAGATGGCGTAGCCCTCATATGGTCTGGGTGGGACAGCAAGCACTAGCAG 1260
1261 GGAAGGATATGATTTATACCGCTGGGGGAATCGATTTCACACCCCACTTCTTTCTCCAC 1320

1261 GGAAGGATATGATTTATACCGCTGGGGGAATCGATTTCACACCCCACTTCTTTCTCCAC 1320
1321 AACAAATTCCTTACCCCTCTAGCCAATGGCGTTTACAAACCATGTTTGGAGCGGCACAGGTC 1380
1321 AACAAATTCCTTACCCCTCTAGCCAATGGCGTTTACAAACCATGTTTGGAGCGGCACAGGTC 1380
1381 CTGTAGATGGCAACGAATGCGACTATCTACCTCCGGGCAAAATGGAACTTGCACCGCATGT 1440
1381 CTGTAGATGGCAACGAATGCGACTATCTACCTCCGGGCAAAATGGAACTTGCACCGCATGT 1440
1441 TCGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTTTGGCAAGGCAATAGTCTCTA 1500
1441 TCGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTTTGGCAAGGCAATAGTCTCTA 1500
1501 GCAAAAAACAACCTTGTAGAACCAAGTAGAAGCGGGCGGATTTGTTTAAATTCATGAAG 1560
1501 GCAAAAAACAACCTTGTAGAACCAAGTAGAAGCGGGCGGATTTGTTTAAATTCATGAAG 1560
1561 ACTGGGCAACAACAAGTGGCATCGATCTCTGCTTGAAGCGGTGGCAGATGAATACGATG 1620
1561 ACTGGGCAACAACAAGTGGCATCGATCTCTGCTTGAAGCGGTGGCAGATGAATACGATG 1620
1621 TGCAAGTTTGTATCCACACCGGATCAAGTCAATGAGCAGGTTATGTAGATGACACCTAA 1680
1621 TGCAAGTTTGTATCCACACCGGATCAAGTCAATGAGCAGGTTATGTAGATGACACCTAA 1680
1681 ATGCAATGAACGGGCGCGCATCCATCCACCAATTCAGGGAGCGGTGGAGGACACT 1740
1681 ATGCAATGAACGGGCGCGCATCCATCCACCAATTCAGGGAGCGGTGGAGGACACT 1740
1741 CACTGTATTTATCAACCATGGCAGGAGCTCAATTTCTACCTCTCCACCAACCCCA 1800
1741 CACTGTATTTATCAACCATGGCAGGAGCTCAATTTCTACCTCTCCACCAACCCCA 1800
1801 CTATTTCCCTATACCAATTAACGTTTGCAGACACTTAGACATGCTCATGACATGCCACC 1860
1801 CTATTTCCCTATACCAATTAACGTTTGCAGACACTTAGACATGCTCATGACATGCCACC 1860
1861 ACCTAGACAAACGCAATCCGCGAGGATTTACAAATTTCTCAAAGCCGTATCCGCCCCGGCT 1920
1861 ACCTAGACAAACGCAATCCGCGAGGATTTACAAATTTCTCAAAGCCGTATCCGCCCCGGCT 1920
1921 CTATCGCGCTCAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGAT 1980
1921 CTATCGCGCTCAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGAT 1980
1981 CGCAAGCAATGGGCGTGCAGGCAAGTATCTCGAACTTGGGAGACTGCGGATGAAGA 2040
1981 CGCAAGCAATGGGCGTGCAGGCAAGTATCTCTCGAACTTGGGAGACTGCGGATGAAGA 2040
2041 ATAAAAAAGAAATTTGGTAACTTCTGAAGATGGCAAGATACGATAATTTCCGCAATTA 2100
2041 ATAAAAAAGAAATTTGGTAACTTCTGAAGATGGCAAGATACGATAATTTCCGCAATTA 2100
2101 AGCGCTACATCTCCAAATACACTATCAACCCGCTTTGACCCACGCGGTGAGGAGTATA 2160
2101 AGCGCTACATCTCCAAATACACTATCAACCCGCTTTGACCCACGCGGTGAGGAGTATA 2160
2161 TCGGCTCTGTGGAAGGCAAGATCGCGACTTGGTGTGTGGAAATCTTCCTTTTGT 2220
2161 TCGGCTCTGTGGAAGGCAAGATCGCGACTTGGTGTGTGGAAATCTTCCTTTTGT 2220
2221 GGTAAAAACCCAAATTCGTATCAAGGCGGTATGGTGTCTCTCTGAATGGCGGAT 2280
2221 GGTAAAAACCCAAATTCGTATCAAGGCGGTATGGTGTCTCTCTGAATGGCGGAT 2280
2281 CTAAAGCGCTGTGCGCACTCCCAACCGGTTTATTAACCGCGAAATGTTTGGGCATCACG 2340
2281 CTAAAGCGCTGTGCGCACTCCCAACCGGTTTATTAACCGCGAAATGTTTGGGCATCACG 2340
2341 GCAAGCGAAATTTGACACCGATCACTTTTGTTCCTCAAGTCGCTATGAAATGCGG 2400

Db 2341 GCAAGCGAAATTGACACACAGCATCACTTTGTTTCCAAAGTCGCTATGAAATGGCG 2400
Qy 2401 TGAAGAAAAGCTGGGCTTAGAGCGCCAGTTCTACCGGTCAAAAATCGCGGTAAACATCA 2460
Db 2401 TGAAGAAAAGCTGGGCTTAGAGCGCCAGTTCTACCGGTCAAAAATCGCGGTAAACATCA 2460
Qy 2461 CCAAGAAAGACTTCAAGTTCAACGACAAAACGCGCAAAAATCACCGTCGATCCGAAAACCT 2520
Db 2461 CCAAGAAAGACTTCAAGTTCAACGACAAAACGCGCAAAAATCACCGTCGATCCGAAAACCT 2520
Qy 2521 TCGAGGTCCTTTGTAGATGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTGCTCTAG 2580
Db 2521 TCGAGGTCCTTTGTAGATGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTGCTCTAG 2580
Qy 2581 CCAGCGCTACATCTTCTTAGGCAAAATGCGCCCTTTGGGGGAGGTTATTTAGGAA 2640
Db 2581 CCAGCGCTACATCTTCTTAGGCAAAATGCGCCCTTTGGGGGAGGTTATTTAGGAA 2640
Qy 2641 TCTTCATCAACACGACCTGCAATCGGTCCTGCGGTGCGATCGTGCTTTAAACAAC 2700
Db 2641 TCTTCATCAACACGACCTGCAATCGGTCCTGCGGTGCGATCGTGCTTTAAACAAC 2700
Qy 2701 TTTTCATCTTTAAGCAATCGCATTTTAAATTAATTTAAATTTCTTATAATTAATATAT 2760
Db 2701 TTTTCATCTTTAAGCAATCGCATTTTAAATTAATTTAAATTTCTTATAATTAATATAT 2760
Qy 2761 TATGCCCTCTCACTTTTAAAGGAGAAATATGCGTAGGTCCTTTGGTATTTGCTATGCGGG 2820
Db 2761 TATGCCCTCTCACTTTTAAAGGAGAAATATGCGTAGGTCCTTTGGTATTTGCTATGCGGG 2820
Qy 2821 TTGTTGGTCTGGGCGCAAGGGTATTGAAACCCATCGCTCTCAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTGGTCTGGGCGCAAGGGTATTGAAACCCATCGCTCTCAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 2
US-09-904-994B-13
; Sequence 13, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(728)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (739)..(2445)
US-09-904-994B-13

Query Match 77.3%; Score 2228; DB 3; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 159 AGGACTATATCCCTTTCCTTTTAAATTTAAACAAGGAGTAAATAGGTGAACTCACAC 218
Db 1 AGGACTATATCCCTTTCCTTTTAAATTTAAACAAGGAGTAAATAGGTGAACTCACAC 60
Qy 219 CCAAGACGAAGAAAAGTTCTTGTATATATCGGGCGGAGTGGCTAGAAAGCGCAAG 278

Db 61 CCAAGACGAAGAAAGTTCTTGTATATATATGCGGCGAAGTGCGCTAGAAACGCGCAAAG 120
Qy 279 CAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTTACATTAGTGCCTATATTATGG 338
Db 121 CAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTTACATTAGTGCCTATATTATGG 180
Qy 339 ACAGAGCGCGCTGGAAAAAAGCCGTTGTCAGCTTATGGAAGAGTGCATGCACCTTTT 398
Db 181 ACAGAGCGCGCTGGAAAAAAGCCGTTGTCAGCTTATGGAAGAGTGTATGCACCTTTT 240
Qy 399 TGAAGAAAGATGAAGTAATGCGCGGCTGGGTAATATGCTTCCCGATCTAGGTGTAGAAG 458
Db 241 TGAAGAAAGACGAGGTGATGCGCGGCTGGGTAATATGCTTCCCTGATTTGGGCGTGAAG 300
Qy 459 CCACCTTTCCGTATGGTACGAAACTTTGTAACTGTGAACTTTGGCCCATCGAACACAGATGAGC 518
Db 301 CCACCTTTCCCGATGGCACCAAACTCGTAACCTGTGAATTTGGCCCATCGAACCTGATGAAC 360
Qy 519 ACTTCAAGCGGCGAAGTGAATTTGGTTGCGATTAAGACATCGAGCTCAATGCAAGGCA 578
Db 361 ACTTCAAGCGGCGTGAAGTGAATTTGGCTGTGATTAAGACATTTGAACCTCAACGCGCAGGA 420
Qy 579 AAGAAGTAACCGAACTCTGAGGTTACTAATGAAGGCGCTAAATCTTGTGATGTGGGTAGCC 638
Db 421 AGGAAGTTACCGAACTAGAGTTACTAACAAGGACCTTAATCTTGTGATGTGGGTAGCC 480
Qy 639 ATTTCCACTTCTTTGAACTAAACAGGCACTAAAAATTCGATCGTGAAGAAAGCTATGGCA 698
Db 481 ATTTCCACTTCTTTGAACTAAACAGGCACTAAAAATTCGATCGTGAAGAAAGCTATGGCA 540
Qy 699 AACGCTTAGATTTCCCTCTGGCAACACGCTACGCAATTTGGGCGCAGGACAAACCCGCAAG 758
Db 541 AACGCTTAGATTTCCCTCTGGCAACACACATACGCAATTTGGGCGCAGGACAAACCCGTAAG 600
Qy 759 TGCAGTTGATTTCTTCTGCTGCGCATTAAGAGTGAATTTGGCATGAACGGGCTTGTGAATA 818
Db 601 TGCAGTTAATCTTCTTGGCGGTAGTAAAAAGTGAATTTGGCATGAACGGGCTTGTGAATA 660
Qy 819 ACATCGGCGATGAACGCCATAAAACATAAAGCGCTTTGACAAAGCGGCAAAATCTACCGGATTTA 878
Db 661 ATATTGCGGACGACGCGCATAAACAAAGCGCTAGACAAAGCAAAATCTACCGGATTTA 720
Qy 879 TCAAGTAAGGAGACTCCCATGAAATGAAACAAAGAAATATGTAAATACCTACGACCC 938
Db 721 TCAAGTAAGGAGACTCCCATGAAATGAAACAAAGAAATATGTAAATACCTACGACCC 780
Qy 939 ACCAAGCGCATAAAGTGGCTTAGGACATCCGATCTTTGGGCGAGAGTAGAACATGAC 998
Db 781 ACCAAGCGCATAAAGTGGCTTAGGACATCCGATCTTTGGGCGAGAGTAGAACATGAC 840
Qy 999 TATACCACTATGCGGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATG 1058
Db 841 TATACCACTATGCGGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATG 900
Qy 1059 GGTGAGACATAGCCCTGTAGTGAACCAACCTTAGATTTAGTTCATCACTAAACCGATGATT 1118
Db 901 GGTGAGACATAGCCCGACATGCAAGTGAACCAACCTTAGATTTAGTTCATCACTAAACCGATGATT 960
Qy 1119 ATGCACTACACCGGATTTTCAAGCGCAGCATTTGGGATTTAAACCGGCAAAATTCATGGC 1178
Db 961 ATTGACTACACCGGATTTTCAAGCGCAGCATTTGGCATTAAAAATTTGGCAAAATTCATGGC 1020
Qy 1179 ATTGGCAAGCGAGGAAACAAAGGACATGCAAGTGGCGTAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGGCAAGCGAGGAAACAAAGGACATGCAAGTGGCGTAAGCCCTCATATGTCGTGGGT 1080
Qy 1239 GTGGGCAACAGACATAGCAGGGGAGGATGATTATTATACCGCTGGGGGAAATCGATTCA 1298
Db 1081 GTGGGCAACAGACATAGCAGGGGAGGATGATTATTATACCGCTGGGGGAAATCGATTCA 1140
Qy 1299 CACACCACTTCTTCTTCTTCCACAAATTCCTACCGCTTAGCGCAATGGCGGTTCAACAC 1358
Db 1141 CACACCACTTCTTCTTCTTCCACAAATTCCTACCGCTTAGCGCAATGGCGGTTCAACAC 1200

```
QY 1359 ATGTTTGGAGGCGGCA CAGGTCTCTGTAGATGGCA CGAATGCGACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGCA CAGGCCCCGTAGATGGCA CGAATGCGACTACTATCACTCCGGGC 1260
QY 1419 AATGGAACCTTGACCCGATGTTGGCGGACGAGAGTATTCTATGAATGGGCTTT 1478
Db 1261 AATGGAACCTTGACCCGATGTTGGCGGACGAGAGTATTCTATGAATGGGCTTT 1320
QY 1479 TTGGCAAAAGCAATAGCTCTAGCAAAAAA CAACTTGTAGAACAGTATAGAGCGGCGC 1538
Db 1321 TTGGCAAAAGCAATAGCTCTAGTAAAAA CAACTTGTAGAACAGTATAGAGCGGCGC 1380
QY 1539 ATTGGTTTTAAATGCAAGACTGGGGCA CAA CCAAGTGCATCGATCACTGCTTG 1598
Db 1381 ATTGGTTTTAAATGCAAGACTGGGGCA CAACTCCAAGTGCATCGATCACTGCTTG 1440
QY 1599 AGCGTGCAGATGATAGATGTCAGTTCAGATGTTGTTATCCACCGATACAGTCAATGAGCA 1658
Db 1441 AGCGTAGCATGATGATAGATGTCAGATGTTGTTATCCACCGATACAGTCAATGAGCA 1500
QY 1659 GGTATATGATGACACCCCTAAATGCAATGAACGGGCGCGCATCCATGCCCTACCACTT 1718
Db 1501 GGTATATGATGACACCCCTAAATGCAATGAACGGGCGCGCATCCATGCCCTACCACTT 1560
QY 1719 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACTGCGAGGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACTGCGAGGCGAGTGAATATT 1620
QY 1779 CTACCCCTCTCCACCCGCCACTATTCCCTATACCACTTAATAGCGTTGCAAGACACTTA 1838
Db 1621 CTACCCCTCTCCACCCGCCACTATTCCCTATACCACTTAATAGCGTTGCAAGACACTTA 1680
QY 1839 GACATGCTCATGACATGCACCATAGACAAACGATCCGCGAGGATTTACAAATTTCT 1898
Db 1681 GACATGCTCATGACATGCACCATAGATTAACGCAATCCGCGAGGATCTCCAATTTCT 1740
QY 1899 CAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGGTG 1958
Db 1741 CAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGGTG 1800
QY 1959 ATCGCGATGACAGCTCGGATTCGACAGCAATGGGCGGTGACGGGAAGTGAATCTCTGA 2018
Db 1801 ATCGCGATGACAGCTTCGATTCGACAGCAATGGGCGGTGACGGGAAGTGAATCTCTGA 1860
QY 2019 ACTTGGCAGCTGGGATAGAAATAAAGAAATTTGGTAGACTTCTCGAAGTGGCA 2078
Db 1861 ACTTGGCAAACTGCAGACAAAGAAATAAAGAAATTTGGTAGACTTCTCGAAGTGGTGA 1920
QY 2079 GATAACGATTAATTTCCGATTAAGCGCTTACATCTCCAAATACATATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAACTTCCGATCAAAAGCTATATCTCCAAATACACATTAATCCCGCTTTG 1980
QY 2139 ACCACCGGTGACGAGTATATCGGCTCTGTGGAAGAGGCAAGATCCCGACTTGTG 2198
Db 1981 ACCATGGGTGACGAGTATATCGGCTCTGTGGAAGAGGCAAGATCCCGACTTGTG 2040
QY 2199 GTGTGGAATCTGCTCTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGGCGGTATG 2258
Db 2041 GTGTGGAATCTGCTCTTTTGGCGTAAACCCCAAAATCGTGATCAAAAGGCGGTATG 2100
QY 2259 GTCTTCTGTAATGGGCGATTTCTAAGCGCTGTGCGCCACTCCCCCAACCGGTTATTAC 2318
Db 2101 GTGTTCTGTAATGGGCGATTTCTAAGCGCTGTGCGCCACTCTCCAGCCGGTTATTAC 2160
QY 2319 CGCGAAATGTTTGGGCACTACGCGAAGCGAAATTTTGA CACGAGCATCACTTTTGTTC 2378
Db 2161 CGCGAAATGTTTGGGCACTACGCGAAGCGAAATTTTGA CACGAGCATCACTTTTGTTC 2220
QY 2379 AAAGTCGCTATGAAATGGGTGAAGAAAGCTGGGCTTAGCGCGCAAGTCTACCG 2438
Db 2221 AAAGTCGCTATGAAATGGGTGAAGAAAGCTTAGGTTTAGAGCGCAAGGTGCTCCC 2280
```

```
QY 2439 GTCAAAAACCTCGCTAA CATTACCAAGAAAGACTTCAAGTTCAA CGACAAAACGGCAAAA 2498
Db 2281 GTGAAAAACCTGCGTTAA CATTACCAAGAAAGACTTCAAGTTCAA CGACAAAACGGCAAAA 2340
QY 2499 ATCACCGTCGATCCGAAAACCTTCAGAGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGTCGATCCGAAAACCTTCAGAGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
QY 2559 CCCACTCGCAAGTGCCTCTAGCCCAAGCGGTACACTTTTCTTAGGCACAAT 2610
Db 2401 CCCACTCTGAAGTGCCTCTAGCCCAAGCGGTACACTTTTCTTAGGCACAAT 2452

RESULT 3
US-09-904-994B-4
; Sequence 4, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2405
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(681)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (692)..(2398)
; US-09-904-994B-4

Query Match 74.9%; Score 2158.6; DB 3; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
```

```
QY 206 GTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATATCGCGGCGAAGTGGCT 265
Db 1 GTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATATCGCGGCGAAGTGGCT 60
QY 266 AGAAGCCCAAGCAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
Db 61 AGAAGCCCAAGCAGAGGGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTACATTAGT 120
QY 326 GCCCATATTATGACCAAGCGCGCTGGAAAAAACCCTTGCCTTATGGAAGAG 385
Db 121 GCCCATATTATGACCGAGGCGCGCTGGCAAAAACCCTTGTCTGNACTTATGGAAGAA 180
QY 386 TGCATGCACTTTTCAAAAAAGATGAAGTAATGCCGGGGTGGTAAATATGTTCCCGAT 445
Db 181 TGTATGCACTTTTCAAAAAAGATGAGGTGATGCCGGTGTGGGGAATATGTTCCCTGAT 240
QY 446 CTAGGTGTAGAGCCACCTTTCTGATGCTAGCAAACTTGTAACTGTGAATTCGGCCATC 505
Db 241 TTGGCGGTAGAAAGCCACTTTCCCGATGGCACAAAACCTCGTAACCGTGAATTCGCCCAT 300
QY 506 GAAACAGATGAGCACTTCAAGCGGGCGAAGTGAATTTGGTTCGATAAAGACATCGAG 565
Db 301 GAACTGATGAACACTTTAAAGCCGGTGAAGTGAATTTGGCTGTGATTAAGACATTGAG 360
QY 566 CTCAATGCGGCAAGAAAGTAAACCGAATTTGAGGTTACTTAATGAAGGGCTAAATCTTG 625
Db 361 CTCAACGCGGGTAAGGAAGTTACCGAGCTTGAAGTTTACCAACGAAGGA CTTAAATCTTG 420
QY 626 CATGTGGGTAGCCATTTCCACTTTCTTTGAGCTTAACAGGCACATAAATTCGATCGTGA 685
Db 421 CATGTGGGTAGCCATTTCCACTTTCTTTGAAACCAACAGGCATTTGAAATTCGATCGGAA 480
```



```

; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(682)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (693)..(2399)
US-09-904-994B-10.

Query Match      74.9%; Score 2158.6; DB 3; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 206 GTGAAAATCTACACCCAAAGAGCAAGAAAGTCTTGTGTTATATATGCGGGCGAAGTGGCT 265
Db 2 GTGAAATCTACACCCAAAGAGCAAGAAAGTCTTGTGTTATATATGCGGGCGAAGTGGCT 61

Qy 266 AGAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAAGCATTTGCTTACATAGT 325
Db 62 AGAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAAGCATTTGCTTACATAGT 121

Qy 326 GCCCATATTATGAGCAGAGCGCGGTGGAAGAAACCGTTGCCAGCTTATGGAAGAG 385
Db 122 GCCCATATTATGAGCAGAGCGCGGTGGAAGAAACCGTTGCCAGCTTATGGAAGAG 181

Qy 386 TGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCGGGTGGGTAAATATGTTCCCGAT 445
Db 182 TGTATGCACCTTTTGAAGAAAGACAGAGTATGCCGGGTGGGTAAATATGTTCCCGAT 241

Qy 446 CTAGGTGTAGAACCCACTTTCTGTAGTGTACGAAACTTTGAACCTGTGAATTTGCCCATC 505
Db 242 TTAGCGTGGAACTACTTTTCCCGATGGCACCACCAACTCGTAAACCGTGAATTTGCCCATC 301

Qy 506 GAACCATATGAGCACTTCAAGCGGGCGAAGTGAATTTGGTTGGGTAAAGACATCGAG 565
Db 302 GAACCCGATGAACCTTCAAGCGGGCGAAGTGAATTTGGTTGGGTAAAGACATCGAA 361

Qy 566 CTCATGCGCAAGAGTAACCGAATCTCAGAGTTACTAATGAAGGCCCTAAATCCTTTG 625
Db 362 CTCAGCGAGTAAGAGTTTACCGAATCTAGAAATTTACCAAGAGGACCTAAATCCTTTG 421

Qy 626 CATGTGGGTAGCCATTTTCCACTTTCTTGAAGCTAACAGGCACTAAAAATTCGATCGTGA 685
Db 422 CATGTGGGTAGCCATTTTCCACTTTCTTGAAGCAACAGAGCATTTGAATTCGATCGGAA 481

Qy 686 AAAGCCTATGCAAAACCGCTTAGATATTCCTCTGTGGCAACACGCTACGCAATTTGGGCGAG 745
Db 482 AAAGCCTATGCAAAACCGCTTAGATATTCCTCTGTGGCAACACGCTACGCAATTTGGGCGAG 541

Qy 746 CAACCCGCAAGTGCAGTTGATTCCTCTGTGGTGCAGTAAAGATGATTTGGCATGAAC 805
Db 542 CAACCCGCAAGTGCAGTTGATTCCTCTGTGGTGCAGTAAAGATGATTTGGCATGAAC 601

Qy 806 GGGCTTGTGAATAACATCGCGGATGAACCGCAATAACATAAGCGCTTTGAACAAGCGGAAA 865
Db 602 GGGCTTGTGAATAATTTGCAAGTGAACCGCAATAACATAAGCGCTTTGAATAAGCAAAA 661

Qy 866 TCTCAGCGATTTTCAAGTAAGAGAGCTCCCATGAAAAATGAAAAACAAGAAATATGTAAA 925
Db 662 TCTCAGCGATTTTCAATAAAGAGAGCTCCCATGAAAAATGAAAAACAAGAGTATGTAAA 721

Qy 926 TACCTACGAGCCCAAGCGGATAAGTGGCTTTAGAGATACCGATCTTTGGGCGAGA 985
Db 722 TACCTACGAGCTTACCAAGCGGCAAAAGTGGCTTTAGAGATACCGATCTTTGGGCGAGA 781

Qy 986 AGTAGAATCATCATATACCATATGCGGAAGAACTTAAATTTGGCGGGGTAAAACTAT 1045
Db 782 AGTAGAATCATCATATACCATATGCGGAAGAGCTCAAAATTTGGCGGGGTAAAACTAT 841

Qy 1046 CCGTGAGGGTATGGGTACAGAGCAATPAGCCCTGATGAAAAACCCCTPAGATTTAGTCATCAC 1105
Db 1046 CCGTGAGGGTATGGGTACAGAGCAATPAGCCCTGATGAAAAACCCCTPAGATTTAGTCATCAC 1105

```

```

Db 842 CCGTGAGGGCATGGGTACAGCAATAGTCCAGATGAATAACCCCTAGATTTAGTCATCAC 901
Qy 1106 TAACCGCATGATTTATCGACTACACCGGGATTTTACAAAGCCGACATTTGGGATTTAAACCG 1165
Db 902 CAACCGCATGATTTATCGACTACACCGGGATTTTACAAAGCCGACATTTGGGATTTAAACCG 961
Qy 1166 CAAATATCCATGGCATTTGGCAAGGAGGAGAAACAAGACATGCAAGATGGGCTTAAGCCCTCA 1225
Db 962 CAAATATCCATGGCATTTGGCAAGGAGGAGAAACAAGACATGCAAGATGGGCTTAAGCCCTCA 1021
Qy 1236 TATGTCGTGGGTGTGGGCACAGAAAGCACTAGCAGGGGAAGGTATGATTTATTTACCGCTGG 1285
Db 1022 TATGTCGTGGGTGTGGGCACAGAAAGCATTAGCAGGGGAAGGTATGATTTATTTACCGCTGG 1081
Qy 1286 GGGAAATCGATTACACACCCACTTCTCTTCTCCCAACAACAAATCCCTACCGCTTAGCCCAA 1345
Db 1082 GGGAAATCGATTACACACCCACTTCTCTTCTCCCAACAACAAATCCCTACCGCTTAGCCCAA 1141
Qy 1346 TGGCGTTACACCATGTTTGGAGGCGGCACAGTCTGTAGATGGCACGAATCGGACTAC 1405
Db 1142 TGGCGTTACACCATGTTTGGAGGCGGCACAGTCTGGGTAGATGGCACGAATCGGACTAC 1201
Qy 1406 TATCACTCCGGCAAAATGGAACCTTGCACCGCATGTTGCCGAGAGAGAGATTTCTAT 1465
Db 1202 CATCACTCCGGCAAAATGGAACCTTGCACCGCATGTTGCCGAGAGATTTCTAT 1261
Qy 1466 GAATGTGGCTTTTGGGCAAAAGGCAATAGCTTAGCAAAAAACCACTTTGTGAACAAAGT 1525
Db 1262 GAATGTGGCTTTTGGGCAAAAGGCAATAGCTTAGCAAAAAACCACTTTGTGAACAAAGT 1321
Qy 1526 AGAAGCGGGCGGATTTGTTTAAATTTGCATGAAGACTGGGGCAACAACCAAGTGGGAT 1585
Db 1322 AGAAGCGGGCGGATTTGTTTAAATTTGCATGAAGACTGGGGCAACAACCTTCAAGTGGCAAT 1381
Qy 1586 CGATCACTGCTTTGAGCGTGGCAGATGAATACCATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1382 CGATCACTGCTTTGAGCGTGGCAGATGAATACCATGTGCAAGTTTGTATCCACACCGATAC 1441
Qy 1646 AGTCAATGAGGCAAGTATGTAGATGACACACCTTAAATGAAGTGAACGGGCGGCCATCCA 1705
Db 1442 GGTCAATGAGGCAAGTATGTAGATGACACACCTTGAATGCGATGAACGGGCGGCCATCCA 1501
Qy 1706 TGCCTACCAATTTGAGGGAGCGGGTGGAGGACACTCACCTGTATTTATCAACATGGCAGG 1765
Db 1502 TGCCTACCAATTTGAGGGAGCGGGTGGAGGACACTCACCTGTATTTATCAACATGGCAGG 1561
Qy 1766 CGAGCTCAATATTTTACCTCTCCACACCGCCCACTATTCCTTATACCATTAATACGCT 1825
Db 1562 CGAGCTCAATATTTTACCTCTCCACACCGCCCACTATTCCTTATACCATTAATACGCT 1621
Qy 1826 TGCAGAACACTTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGA 1885
Db 1622 TGCAGAACACTTTAGACATGCTCATGACCTGCCACCACTAGATATAACGATCCGCGAGGA 1681
Qy 1886 TTTTCAATTTTCTCAAAAGCGGTATCGCCCGGCTCTATCGGGCTGGAAGATGTCTCCA 1945
Db 1682 TTTTCAATTTTCTCAAAAGCGGTATCGCCCGGCTCTATCGGGCTGGAAGATGTCTCCA 1741
Qy 1946 TGATATGGGTGTGATCGCGATGACAGCTCGGATTCGACGAATGGGCGTGCAGCGGA 2005
Db 1742 TGATATGGCGTGTGATCGCGATGACAGCTCGGATTCGACGAATGGGCGTGCAGCGGA 1801
Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATTAAGATAAAAAAGAAATTTGGTAAAGCTTCC 2065
Db 1802 AGTGATTTCTCGAACTTTGGCAGACTGCGGATTAAGATAAAAAAGAAATTTGGTAAAGCTTCC 1861
Qy 2066 TGAAGATGSCAAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1862 TGAAGATGATGCGAGATAACGATACTTCCGTATCAAAACGCTACATCTCCAAATACACTAT 1921
Qy 2126 CAAACCCCGCTTTGACCCACGGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 2185
Db 1922 TAAACCCCGCTTTAAACCCATGGGGTAAAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAAAT 1981

```

QY 2186 CGCGACTTGGTGGTGGATCTCTGCTTTTGGCGCTAAACCCAAATCGTGATCAA 2245
DB |||||
QY 1982 CGCTGATTGGTGGTGGATCTCTGCTTTTGGCGCTAAACCCAAATCGTGATCAA 2041
DB |||||
QY 2246 AGCGGATATGGTGGTGGTCTCTGAAATGGCGGATCTAAACCGCTCTGGCCCACTCCCA 2305
DB |||||
QY 2042 AGCGGATATGGTGGTGGTCTCTGAAATGGCGGATCTAAACCGCTCTGGCCCACTCA 2101
DB |||||
QY 2306 ACCGGTTATTAACCGCGAAATGTTTGGGATCATCGGCAAGCGAAATTTGACACCAAGAT 2365
DB |||||
QY 2102 GCGGTTTATTAACCGCGAAATGTTTGGGATCATCGGCAAGCGAAATTTGACACCAAGAT 2161
DB |||||
QY 2366 CACTTTTGTTCCTAAAGTCGCTATGAAATGGCGGATGAAAGAAAGCTGGGCTTAGAGCG 2425
DB |||||
QY 2162 CACTTTTGTTCCTAAAGTCGCTATGAAATGGCGGATGAAAGAAAGCTTAGAGCG 2221
DB |||||
QY 2426 CCAAGTTCTACCGGTCAAAACTCGCGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGA 2485
DB |||||
QY 2222 CAAGGTGCTACCGGTGAAACTCGCGCAACATCACTAAGAAAGACTTCAAAATCAACAA 2281
DB |||||
QY 2486 CAAACCGCAAAATCAACGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAACT 2545
DB |||||
QY 2282 CAAGCGGCGATATCACTGTCGATCTTAAACCTTCGAGGTCTTTGTAGATGGCAACT 2341
DB |||||
QY 2546 CTGACCTCTAAACCCCACTCGCAAGTGCCTTAGCCAGCGCTACACTTTCTTAGGC 2605
DB |||||
QY 2342 CTGACCTCTAAACCCCGCTCTGAAGTGCCTTAGCCAGCGCTACACTTTCTTAGGC 2401
DB |||||
QY 2606 ACAATG 2611
DB |||||
QY 2402 NCAATG 2407
DB |||||

RESULT 5

US-09-904-994B-7
; Sequence 7, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degosedesquences
; CURRENT APPLICATION NUMBER: US/09/904, 994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2183
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(683)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (694)..(2181)
US-09-904-994B-7

Query Match 67.6%; Score 1948; DB 3; Length 2183;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 206 GTGAACTCACCCCAAGAGCAAGAAAGTCTTGTTATATATATGCGGAGTGGCT 265
DB |||||
QY 3 GTGAACTCACCCCAAGAGCAAGAAAGTCTTGTTATATATGCGGAGTGGCT 62
DB |||||
QY 266 AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCCAAGCAATGCTTACATTAGT 325
DB |||||
QY 63 AGAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCCCAAGCAATGCTTACATTAGT 122
DB |||||
QY 326 GCCATATTATGACGAGCGCGCTGGAAAAAACCCTGTGCCAGCAAGAGTATTTCTAT 385
DB |||||
QY 123 GCCATATTATGACGAGGCGCGCTGGCAAAAAACCCTGTGCTGAACCTTATGGAAGAA 182
DB |||||

QY 386 TGCATGCACCTTTTGAAGAGATGAATATGCGCGGGTGGGTAAATATGTTCCGAT 445
DB |||||
QY 183 TGTATGCACCTTTTGAAGAGATGAATATGCGCGGGTGGGTAAATATGTTCCGAT 242
DB |||||
QY 446 CTAGGTGTAGAAGCCACCTTTCTGATGTTGACCAAACTTCTAACTGTGAATTTGGCCCATC 505
DB |||||
QY 243 TTGGGGGTAGAAGCCACCTTTCCCGATGGCAAACTGTAACCGGTGAATTTGGCCCAT 302
DB |||||
QY 506 GAACAGATGAGCAGCTTCAAGCGGCGGAGTGAATTTGGTTCGATAAAGACATCGAG 565
DB |||||
QY 303 GAACCTGATGAACACCTTTAAAGCGGTGAAGTGAATTTGGCTGTGATAAAGACATTTGAG 362
DB |||||
QY 566 CTCAATGCAAGGCAAGAGTAACCGAATTTGAGGTACTTAATGAAGGCGCTAAATCTTTG 625
DB |||||
QY 363 CTCAAGCTGGGTGAAGGAGTTTACCGAGCTTGAAGTTTACCAACGAAGGAGCTAAATCTTTG 422
DB |||||
QY 626 CATGTGGGTGAGCAGCTTTCCACTTCTTTGAAGCTTAACAAAGSCACTAAATTTGATCTGTA 685
DB |||||
QY 423 CATGTGGGTGAGCAGCTTTCCACTTCTTTGAACCAACAAAGGCAATGAATTTGATCGGAA 482
DB |||||
QY 686 AAAGCTTATGCAAAACGCTTAGATATTCCTCTCTGGCAACACGCTAGCGATTGGGCGAGGA 745
DB |||||
QY 483 AAAGCTTATGCAAAACGCTTAGATATTCCTCTCTGGCAACACGCTAGCGATTGGGCGAGGA 542
DB |||||
QY 746 CAAACCCGCAAGTGCAGTTGATTTCTTGTGTGGCAGTAAAGTGAATTTGGCATGAAC 805
DB |||||
QY 543 CAAACCCGTAAGTGCAGTTAATTCCTCTTGGCGGTAGTAAAGTGAATTTGGCATGAAC 602
DB |||||
QY 806 GGGCTTGTGAATAACATCGCGATGAAACGCTTAACATTAAGCGCTTGAACGAGGAAA 865
DB |||||
QY 603 GGGCTTGTGAATAATTTGCGGACGAAACGCTTAACCAAGGCACTAGAACAGGCAAAA 662
DB |||||
QY 866 TCTCAGGATTTTCAAGTAAGGAGACTCCCTCAAGTAAAGTGAATTAAGTAAATATGTA 925
DB |||||
QY 663 TCTCAGGATTTTCAAGTAAGGAGACTCCCTCAAGTAAAGTGAATTAAGTAAATATGTA 722
DB |||||
QY 926 TACTTACGGAACCCACCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGA 985
DB |||||
QY 723 CACTTACGGAACCCACCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGA 782
DB |||||
QY 986 AGTAGAATGACTATACCACTATGCGGAAGAACTTAAATTTGGCGGCGGTAAACTAT 1045
DB |||||
QY 783 AGTAGAATGACTATACCACTATGCGGAAGAGCTCAAAATTTGGCGGCGGTAAACTAT 842
DB |||||
QY 1046 CCGTGAGGATGGGTGAGGCAATAGCCCTCATGAACACCCCTAGATTTAGTCATCAC 1105
DB |||||
QY 843 CCGTGAGGATGGGTGAGGCAATAGCCCAATAGCCCAATGAACACCTTAGATTTAGTGATCAC 902
DB |||||
QY 1106 TAAACGATGATTTATGACTACACCGGATTTTCAAGCGGACATTTGGGATTTAAAGAACGG 1165
DB |||||
QY 903 CAACGATGATTTATGACTACACCGGATTTTCAAGCGGACATTTGGTATTTAAAGATGG 962
DB |||||
QY 1166 CAAATCCATGGCATTTGCAAGGCGAGGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCA 1225
DB |||||
QY 963 CAAATCCATGGTATTTGCAAGGCGGGAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1022
DB |||||
QY 1226 TATGTCGTGGTGGGCAAGAGCACTAGCAGGGGAGGTATGATTTATACCGCTCG 1285
DB |||||
QY 1023 TATGTCGTGGTGGGCAAGAGCACTAGCAGGGGAGGTATGATTTATACCGCTCG 1082
DB |||||
QY 1286 GGGATCGATTTACACACCCCTTCTTCTCCCAAACTTCCCTACCGCTTACGCCAA 1345
DB |||||
QY 1083 GGGATCGATTTGCAACCCCTTCTTCTCCCAAACTTCCCTACCGCTTACGCCAA 1142
DB |||||
QY 1346 TGGGCTTACAAACCATTTTGGAGGCGGCAAGGTCTCTGTAGATGGCAGCAAGATGCGACTAC 1405
DB |||||
QY 1143 TGGTGTACAAACCATTTTGGAGGCGGCAAGGTCTCGGTAGATGGCAGCAAGATGCGACCAC 1202
DB |||||
QY 1406 TATCACTCCGGGCAAAATGGAACCTTGCAACCGCATGTTGGCGCAGCAGAGATTTCTAT 1465
DB |||||
QY 1203 CATCACTCCGGGCAAAATGGAACCTTGCAACCGCATGTTGGCGCAGCTGGAAGATTTCTAT 1262
DB |||||

Qy	1466	GAATGTGGGCTTTTGGGCAAGCGCAATAGCTCTGCAAAAAACAACCTTGTAGACAAGT	1525
Db	1263	GAATGTAGGCTTTTGGGCAAGCGCAATAGCTTCTAGCAAAAAACAACCTTGTAGACAAGT	1322
Qy	1526	AGAAGCGGCGCGATTCGTTTAAATTTGCATGAAGACTGGGGCACAAACCAAGTCGCGAT	1585
Db	1323	AGAAGCGGCGCGATTCGCTTTAAATTTGCATGAAGACTGGGGCACAAACCAAGTCGCGAT	1382
Qy	1586	CGATCACTGCTTGAGCGGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGCATAC	1645
Db	1383	CGATCACTGCTTGAGCGGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGCATAC	1442
Qy	1646	AGTCAATGAGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGCGCATCCA	1705
Db	1443	GGTCAATGAGCGAGGTTATGTGGATGACACCCCTAAATGCAATGAACGGGGCGCGCATCCA	1502
Qy	1706	TGCCCTACCACATTCAGGCGAGCGGTGGAGAGCACTCACCTGATGTTATCACCATGGCAGG	1765
Db	1503	TGCCCTACCACATTCAGGCGAGCGGTGGAGAGCACTCACCTGATGTTATCACCATGGCAGG	1562
Qy	1766	CGAGCTCAATATTTCTACCCCTCTCCACACCCGCCACTATTTCCCTATACCAATTAAATCGGT	1825
Db	1563	CGAGCTCAATATTTCTACCCCTCTCCACACCCGCCACTATTTCCCTATACCAATTAAATCGGT	1622
Qy	1826	TGCAGAACACTTAGACATGCTCATGACATGCGACACCTGACAGAACCGCATCCGCGAGGA	1885
Db	1623	TGCAGAACACTTAGACATGCTCATGACATGCGACACCTGACAGAACCGCATCCGCGAGGA	1682
Qy	1886	TTTACAAATTTTTCTCAAAGCGGTATCCGCCCGGCTCTATCCGGCTGAAGATGTCTCCA	1945
Db	1683	TTTACAAATTTTTCTCAAAGCGGTATCCGCCCGGATCTATTTGCCGTGAGAGATGTCTCCA	1742
Qy	1946	TGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGSCGA	2005
Db	1743	TGATATTTGGGTGTGATCGCGATGACTAGCTCCGATTCGCAAGCAATGGGGCGCTGGGGA	1802
Qy	2006	AGTGAATTCCTCGAACTTTGGCGNAGCTGGGATAGAAATAAAAAAGAAATTTGGTAAGCTTCC	2065
Db	1803	AGTGAATTCCTAGAACTTTGGCAAACTGCAGATAGAAATAAAAAAGAAATTTGGTAAGCTTCC	1862
Qy	2066	TGAAGATGGCAAGATACCGATAATTTCCGATTTAAGCGGTACATCTCCAAATACACTAT	2125
Db	1863	TGAAGATGGTGCGATACGAACTTCCGATCCTAAAGCGTATATCTTCCAAATACACCAT	1922
Qy	2126	CAACCCCGCTTTTGACCCAACCGCGTGAAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT	2185
Db	1923	TAAATCCGCTTTTGACCCAACCGCGTGAAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT	1982
Qy	2186	CGCCGACTTGGTGTGTGGAAATCTGCGCTTTTGGCGTAAACCCCAAAATTCGTGATCAA	2245
Db	1983	CGCCGACTTGGTGTGTGGAAATCTGCGCTTTTGGCGTGAACCTCAAGATTTGTGATTA	2042
Qy	2246	AGGCGGTATGGTGTGCTTCTCTGAATGGGCGATTCCTAAACGCTGTGTGCCCACTCCCCA	2305
Db	2043	AGGTGGCAATGGTGTGCTTCTCTGAATGGGCGATTCCTAAACGCTGTGTGCCCACTCCCA	2102
Qy	2306	ACCGGTTTATTTACCGCAAAATGTTTGGGCGATCAGCGCAAGCGCAAAATTTGACACCCAGCAT	2365
Db	2103	GCCCGTTTATTTACCGCAAAATGTTTGGGCGATTCCTAAACGCTGTGTGCCCACTCCCA	2162
Qy	2366	CACTTTGTGTTTCCAAAGTCG	2385
Db	2163	CACTTTGTGTTTCAAGCG	2182

RESULT 6
US-10-639-273-1
; Sequence 1, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian

; TITLE OF INVENTION: HELICOBACTER BIZOZERONII UREASE GENES AND THEIR USES									
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS									
; FILE REFERENCE: 19603/3881									
; CURRENT APPLICATION NUMBER: US/10/639,273									
; CURRENT FILING DATE: 2003-08-12									
; PRIOR APPLICATION NUMBER: 60/404,337									
; PRIOR FILING DATE: 2002-08-16									
; NUMBER OF SEQ ID NOS: 41									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 1									
; LENGTH: 8407									
; TYPE: DNA									
; ORGANISM: Helicobacter bizozeronii									
US-10-639-273--1									
Query Match	39.4%;	Score 1134.4;	DB 8;	Length 8407;					
Best Local Similarity	68.5%;	Pred. No. 1.7e-238;							
Matches 1681;	Conservative	0;	Mismatches 741;	Indels 32;	Gaps 2;				
Qy	183	AAATTTAACACAAGGAGTAAATAGGTGAACTCACCCCAAAGAGCAAGAAAGTTCTTGT	2424						
Db	2319	AAATTTGGTGAAGAGAGTTTAGGATGAAATTAACCCCTAAAGAGCTGGCAAGCTCATGT	2378						
Qy	243	TATATTATCGGGCGAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAAC	302						
Db	2379	TGCATTATCGGGCGAATTGGCTAAAAAAGCGCAAGCAAAATGCGTTAAAGCTAAATTTATA	2438						
Qy	303	CCGAAGCCATTGCTTACATTAGTGGCCCATATTATGCAAGAGCGCCGCTGGAAAAA	362						
Db	2439	CTGAGGCGTAGCCCTCATCTGTCCTCATGTGATGGAAGAGCCGCTGCAGGTAAAAA	2498						
Qy	363	CCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGTGAAAAAAGATGAAGTAATGCCCG	422						
Db	2499	GTGTGGCGGATTGTGATGCAAGAGGCGAGGACACTTCTTAAAGCTGATGATGTCTATGCCCG	2558						
Qy	423	GGGTGGGTATATGGTTCGGATCTAGGTGTAGAAGCCACCTTCTCTGATGGTACGAAC	482						
Db	2559	GTGTAGGCCCATATGATCCACGAAGTGGGGAATTGAAGCTAACTTCCCTGATGGGACAAAC	2618						
Qy	483	TTGTAACTGTGGAATTGGCCCATCGAACCAAGTACGACACTTCAAAGCGGGCGAAGTGAAT	542						
Db	2619	TGTTAACCATCCATACCCCGTTGAAGATGGTGGGCATAAATTTGGCTCCGGTGAAGTGA	2678						
Qy	543	TTGGTTGCGATAAAGACATCGAGCTCAATGACGGCAAGAAAGTAACCGAACTTTGAGGTTA	602						
Db	2679	TTTTGAAAAACGAAGACATCACTTTGAATGACGAGCAAAACAAAGCCACCATTTTGAAGTGC	2738						
Qy	603	CTAATGAGGCGCTAAATCTTTCGATGGGTAGCCATTTCCACTTCTTTTGAAGCTTAACA	662						
Db	2739	ATAACAAAGCGGATCGCCCCGTGCAGTGGGGCTCCCACTTCCACTTCTTTTGAAGTGAATA	2798						
Qy	663	AGGCACCTAAATTCGATCGTGAATAAAGCCCTATGGCAAAACGCTTAGATATTTCCCTCTGGCA	722						
Db	2799	AGCTTTTGGAAATTTGATCGTGAATAAAGCCCTATGGCAAAACGCTTAGATATTTCTTGGAA	2858						
Qy	723	ACAGCTACGCATTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTTCTTTGGTGGCA	782						
Db	2859	CCGCTGTGCGCTTTGAACCCCGGTGAGAAAAAACCCTGGAAATTTGATTTCAAATTTGGCGTA	2918						
Qy	783	GTAATAAGTGTATGGCATGAACGGGCTTTGTGAATAACATCCGGGTGAACGCCATAAC	842						
Db	2919	ACCAACGCATTTAGCGCTTTAACTCTCTTGTGTGATCGCAAGCCGATCTGATGGCAAAA	2978						
Qy	843	ATAAAGCGCTTGACAAGCGGAATCTCACGGATTT-----	877						
Db	2979	AACTTGTCTCTAAACGCCCAAGAACATCGCTTTTGGTGTGTGAATTCGGTTGCGATA	3038						
Qy	878	-----ATCAAGTAAGGAGACTCCCATG---AAAAATGAAAAAACCAAGATAATGTAAATACCT	930						
Db	3039	AAAAAATAGGAAGGACAAATCCGATGAAAAAATAATCTCGAAAAAGATAATGTTTCTATGT	3098						
Qy	931	ACGAGCCACCAAGGCGGATTAAGTGGGCTTAGGAGATACCGATCTTTGGGCAAGTAG	990						

Db 3099 ATGACCCCACTACCGGCGATAAAGTGAGATTGGCGATACCGACCTGATCTTAGAAGTCG 3158
Qy 991 AACATGACTATACCACTATCGCGAGCACTTAATTTTGGCGGGTAAACATCTCCGTG 1050
Db 3159 AACATGACTGACCACTTATGGCGAAGAAATTAAGTTTGGTGGCGTAAACCACTTCGCG 3218
Qy 1051 AGGTATGGGTGACAGCAATAGCCCTGATGAAACACACCTAGATTAGTCTATCACTAACG 1110
Db 3219 ATGGATGGCAAAACCAAGCCCGACGACCGCAACTCGAICTTGTGCTCACTAACG 3278
Qy 1111 CGATGATTATGACTACACCGGGATTTACAAGCCGACATTTGGGATTAATAACCGCAAAA 1170
Db 3279 CCCTGATCGTGGATTAACACCGGCATTTATAAGCCGATTTTGGCATTAATAAATGCAAAA 3338
Qy 1171 TCCATGTCATTTGGCAAGCAGGAAACAAAGCAATGCAAGATGGGTAGCCCTCATATGG 1230
Db 3339 TCCATGTCATTTGGCAAGCAGGCAATTAAGCAATGCAAGATGGGTGTCACCAATCTTT 3398
Qy 1231 TCGTGGGTGGGCACAGAGCACTAGCAGGGAGGTATGATTTATACCGCTGGGGAA 1290
Db 3399 GCGTGGGCCCTGCTACTGAGGCTTTGGCCGCTGAAGGCTGATTTGTACAGCTGGTGGGA 3458
Qy 1291 TCGATTACACACCACTTCTCTCCACAACAAATCCCTACCGCTCTAGCCCAATGGCG 1350
Db 3459 TTGACACCCACATCCACTTTATTTCTCCCAACAAATCCCAAGCATTTGCCAGCGGA 3518
Qy 1351 TTACAACATGTTTGGAGCGGCAAGGCTCTGATGATGGCAAGATGCGACTACTATCA 1410
Db 3519 TCACAACCATGATTTGGTGGGGAACAGGTCCAGCTGATGGGACTAAGCGGACTACCATCA 3578
Qy 1411 CTCGGGCAATGGAACTTGCACCGCATGTTGCGCGCAGCAGAGAGATTTCTATGAATG 1470
Db 3579 CTCGGGCGGCTGGAACTTTAAACCAATGCTCCGTGCTCTGAAGAATATGCCATGAAT 3638
Qy 1471 TGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAAAGTAGAAG 1530
Db 3639 TGGGCTATTTGGTAAAGGAATGTGTTATGAACCCCTCCCTGGTGCATCACTCGAAG 3698
Qy 1531 CGGGCGGATTTGTTTAAATTTGCATGAAGACTGGGGCAACAACCAAGTCGCGATCGATC 1590
Db 3699 CTGGAGCAATTTGCTTTAAATCCACGAAGACTGGGTAGCACACCTTGCAGCCATCTACC 3758
Qy 1591 ACTGCTTGAGCGTGGCAGATGAATACAGTGTGAAGTTTGTATCCACCGCATACAGTCA 1650
Db 3759 ATTGCTTGAATGTGGTGCACAAATACGATGTGCAAGTGGCTATCCACACCGATACCTTGA 3818
Qy 1651 ATGAGCAGGTTATGTAGATGACACCTTAATGCAATGAACGGCGCGCCATCCATCCCT 1710
Db 3819 ATGAAGCGGCTGTGTGAAGACACTTTTGAAGCCATTTGTCGGCGCACTATCCACACTT 3878
Qy 1711 ACCACATTGAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGGCAGCGCAGC 1770
Db 3879 TCCACACTGAAGTGTGTGGCGGCAACGCTCGGATGCTTAAGATGTCTGGCGAAT 3938
Qy 1771 TCAATATTCTACCTCTCCACACCCCACTATTCCTATATCCATTAATACGGTTGCGAG 1830
Db 3939 TTAACATCTCCAGCTTCTACCAACCCCACTTCTCTTCCAGCTGAATACAGACCG 3998
Qy 1831 AACACTTAGACATGCTCATGACATGCAACACCTAGCAACCGATCCGCGAGATTTAC 1890
Db 3999 AACACATGGACATGTTGATGGTGTGCCACCACTTTGGATATAAAACATCAAAAGATGTC 4058
Qy 1891 AATTTTCTCAAGCCGTATCCGCGCGCTCTATCGCGGCTGAAGATGTCCTCATGATA 1950
Db 4059 AGTTTCTGATTTAGATTTCGCCCCCAACCACTTCGCCGCTGAGGACAAATCCACGATA 4118
Qy 1951 TGGGTGTGATTCGCGATGCAAGCTCGGATTCGAAGCAATGGGCGCTGCGAGCGGAAGTGA 2010
Db 4119 TGGGATTTCTCTATCACAGCTCTGACTCCCAAGCGATGGGCCGTGTAGCGAGGTCA 4178
Qy 2011 TTCCTCGAATTTGGCAGACTGCGGATGAAGTAATAAAGAAATTTGGTAAGCTTCTGAAG 2070
Db 4179 TCACCCGCACTTTGGCAACAGCGCGCAAAAAAATAAAGAAATTTGGTTCGCTTCGCTGAGG 4238

Qy 2071 ATGGCAAGATAACGATAATTTCCGCATTAAGCGCTACATCTCCAATACACTATCAACC 2130
Db 4239 AAAAAGCGGATAATGACAACTTCCGCATCAAGCGCTACATTTCAAATACACCATCAACC 4298
Qy 2131 CGGCTTTGACCCACCGCGGTGAGCGAGTATCGGCTCTGTGGAAGAGGGCAAGATCGCG 2190
Db 4299 CGCTATTGACACCGGCACTTCTGAATATGTGGCTCTGTAGAAAGTGGGCAATTCGCG 4358
Qy 2191 ACTTGGTGGTGGAAATCTGCTTTTGGCGTAAACCCAAATCTGTATCAAGCGG 2250
Db 4359 ATTTGGTGGTCTTCTGAAATGGGCGATTTAAACCGCATTAACCCCAATCATCAAGCGG 4418
Qy 2251 GTATGGTGGTCTTCTGAAATGGGCGATTTAAACCGCTGTGCGCATCTCCCAACCGG 2310
Db 4419 GATTCATCGCACTTTCTCAAATGGGCGATGCCAATGCTCTATCCCACTCCCAACCGG 4478
Qy 2311 TTATTTACCGCAAAATGTTTGGGCGATCAGCGAAGCGGAAATTTGACACGAGCATCACTT 2370
Db 4479 TGTATTACCGCAAAATGTTTGGCCACCATGGTAAAGCCAAATTTGACCAATATCACTT 4538
Qy 2371 TTGTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCCAAG 2430
Db 4539 TTGATCCCAAGTGGCTTATGACACGGCATTTAAAGAGAGTTGGGCTTGCAAGAGTGG 4598
Qy 2431 TTCTACCGGTCAAAAACCTGCGTAACATCAACAAGAAAGACTTCAAGTTCAACGACAAA 2490
Db 4599 TTTTCCAGTTTAAAACTGCGCAACATCACCAAAAGAGCTCAAAATTCACGATGTTA 4658
Qy 2491 CGGCAAAATCACCGTCGATCCGAAACCTTCAGAGTCTTTGTAGATGGCAAACTCTGCA 2550
Db 4659 CGGCACATCGAAGTCAATCTGAAACCTTAAAGTTAAAGTGGATGGCAAGAGGTTA 4718
Qy 2551 CCTTAAACCCCACTCGCAAGTGCCTTAGCCCGAGCTACACTTTCTTCTAGG 2604
Db 4719 CTTCCAAAGCAGCGGATAAATCAGCCTAGCACAACTCTACACTTGTCTAGG 4772

RESULT 7

US-10-639-273-4
; Sequence 4, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-4

Query Match 32.7%; Score 940.8; DB 8; Length 1710;
Best Local Similarity 72.0%; Pred. No. 3e-196;
Matches 1227; Conservative 0; Mismatches 477; Indels 0; Gaps 0;
Qy 900 AAAATGAAAACCAAGATATGTAATACCTACGACCCCAACAAAGCGGATAAAGTCGC 959
Db 7 AAATCTCTCGAAAAGAAATATGTTTCTATGTATGACCCACTACGGCGGATAAAGTGAGA 66
Qy 960 TTAGGATACCGATCTTTGGCGAGTAGAATAGATATACCATCTATCCACCTATCGGAGAA 1019
Db 67 TTGGCGGATACCGACCTGATCTTAGAAGTCGAACATGACTGCACCACTTATGCGGAGAA 126

Qy	1020	CTTAAATTTGGCGGGGTAAAACTATCCGTTGAGGGTATGGGTGAGAGCAATAGCCCTGAT	1079
Db	127	ATTAAAGTTTGGTGGCGGTAAAAACAATTCGCGATGGGATGGCACAACAAACGACCCCGAG	186
Qy	1080	GAACAACACCTAGATTTAGTCATCACCTAACGCGATGATTTACACTACACCCGGATTTAC	1139
Db	187	AGCCAGAACTCGATCTTGTGCTCACTAACCGCCCTGATCGTGAATTTACACCGGCATTTAT	246
Qy	1140	AAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCAGGAACAAG	1199
Db	247	AAAGCCGATTTGGCATTTAAAAATGCAAAATCCATGGCATTTGGCAAGGCAGCAATAAA	306
Qy	1200	GACATGCAAGATGGCGTAAGCCCTCATATGGTCGTGGGTGTTGGGCACAGAGCACTAGCA	1259
Db	307	GACATGCAAGATGGCGTTTGCACAAATCTTTGGCTGGGCCCTCTACTAGGCTTTGGGC	366
Qy	1260	GGGGAAGGTATGATTTATACCGCTGGGGGAATCGATTTACACACCCACACTTCTTTCTCCA	1319
Db	367	GCTGAAGGGCTGATTTGTACAGCTGTGGGATTTGACACCCACATCCACTTTATTTCTCCG	426
Qy	1320	CAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTTACAACCATGTTTGGAGCGGCACAGGT	1379
Db	427	CAACAATCCCACAGCAATTTGCCAGCGGATCACAAACATGATTTGGTGGGGAAACAGGT	486
Qy	1380	CTGTAGATGGCAGGAATGCGACTATCACTCCGGGCAAAATGGAACTTTGCACCCGATG	1439
Db	487	CCAGCTGATGGGACTAACCGGCACTACCATCACTCCGGGGCGCTGGAAACCTTTAAACCATG	546
Qy	1440	TTGGCGCACAGAGAGTATCTNTGAATGTGGGCTTTTGGGCAAAAGCAATAGCTCT	1499
Db	547	CTCGTGCCCTCTGAAGAATATGCCATGAACCTTGGGCTATTTGGGTAAAGGAATGTGCT	606
Qy	1500	AGCAAAAAACAATTGTAGAACAAATGAGAAGCGGGCGCGATTTGTTTAAATTTGCATGAA	1559
Db	607	TATGAACCCCTCCCTGTCGATCAACTCGAAGCTGGAGCAATGGCTTTAAATCCAGAA	666
Qy	1560	GACTGGGGCAACAACAAGTGGATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT	1619
Db	667	GACTGGGGTAGCACACCCTGCAGCCATCTACCAITTGTTGAATGTGGCTGACAAATACGAT	726
Qy	1620	GTGCAAGTTTGTATCCACACGGATACAGTCAATGAGGCGAGTTATGTAGATGACACCCCTA	1679
Db	727	GTGCAAGTGGCTATCCACACCGATACCTTGAATGAAGCGGGCTGTGTGAAGACACTTTG	786
Qy	1680	AATGCAATGAACGGCGCGCCATCCATGCCCTACACATTCAGGAGCGGGTGAGAGACAC	1739
Db	787	CAAGCCATTTGCTGGGCGCATATCCACACTTTCCACACTGAAGGTGTGTGGCGGGCAC	846
Qy	1740	TCACCTGATGTTATCACCATGGCAGGCGAGCTCAATATTTACCCCTCTCCACCAACCCCC	1799
Db	847	GCTCCGATGTCAATTAAGATGTCTGGCGAATTTAAACATCTCCAGCTTCTACCAACCCC	906
Qy	1800	ACTATTCCTTATACCATTAATACGGTTGCAGAACACTTTAGACATGCTCATGACATGCCAC	1859
Db	907	ACCATTCCTTTACCGGTAATACAGAAGCCGACACATGACACATTTGATGGTGTGCCAC	966
Qy	1860	CACCTAGACAAAACCATCCCGGAGATTTACAAATTTCTCAAAAGCGTATCCGCCCGGCG	1919
Db	967	CACTTGATTAATAAACATCAAGAAGATGTCCAGTTTCTGATTTCTAGGATTCGCCCCCAA	1026
Qy	1920	TCTATCGCGGTGAAGATGTGCTCCATGATATGGGTGTATCGCGATGACAAAGCTCGGAT	1979
Db	1027	ACCATCGCGGTGAGGACAAACTCCACGATATGGGATTTTCTCTATCACAGCTCTGAC	1086
Qy	1980	TCGCAAGCAATGGGGCGTGCAGGCGAAGTGAATTCCTCGAACTTGGCAGACTGGCGATAG	2039
Db	1087	TCCCAAGCGATGGGCGGTGTAGGCGAGGTCAATCACCCGCACTTTGGCAAAACAGCGGACAA	1146
Qy	2040	AATAAAAAAGAAATTTGTTAAGCTTCTTGAAGATGGCAAGATAACGATAATTTCCCGATT	2099
Db	1147	AACAAAAAGAAATTTGTTGCTGCTGAGGAAAAAGCGGATATGACACTTTCCGATC	1206
Qy	2100	AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACCCACGCGGTGAGCGATAT	2159

[illegible]

RESIST. T. 8

```

US-10-476-313-10
; Sequence 10, Application US/10476313
; Publication No. US20040241175A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, BRENT
; APPLICANT: BOYLE, JEFFREY
; APPLICANT: LEW, ANDREW
; TITLE OF INVENTION: Antigen Targeting
; FILE REFERENCE: BDWP-002
; CURRENT APPLICATION NUMBER: US/10/476,313
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PR5241
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Helicobacter felis
US-10-476-313-10

```

	Query Match	31.3%	Score	902.4	DB	9	Length	1719	
	Best Local Similarity	70.6%	Pred. No.	8.7e-188					
	Matches 1203	Conservative	0	Mismatches	501	Indels	0	Gaps	0
QY	900	AAAATGAAAAAACAAGAAATATGTAAATACCTACGAGACCCCAAGCGCATAAAGTGC	959						
Db	15	AAGATTTCACGAAGAATATGTTCTATGTATGCTCCACTACCGGGATCGTGTAGA	74						
QY	960	TTAGGAGATACCGATCTTTGGCGAGAAGTAGAACATGACTATACCACTATGGCGAAGAA	1019						
Db	75	CTCGCGGACACTGATTTTGATCTTAAAGTGGAGCATGATTGCACCACTTATGTGTGAAGAG	134						
QY	1020	CTTAAATTTGGCGCGGTAAACACTCCGTGAGGTATGGTTCAGACCAATAGCCCTCAT	1079						
Db	135	ATCAAAATTTGGGGCGGTAAACATCTCCGTGATGGATGAGTCAAAACCAATAGCCCTTAC	194						

Qy		1080	GAAACACCCTAGATTTTAGTCACTAACCGCGATGATTATPCGACTACACCGGGAATTTC	1139
Db		195	TCTTATGAATTAGATTGGTGCTCAGCTAAGCCGCCCTCATGTGGACTATACGGGCATTTAC	254
Qy		1140	AAGCCGACACATTCGGGATTTAAAACGGCABAATCCATGGCATTCGCAAGGCAGGAACAAG	1199
Db		255	AAAGCCGACATTCGGGATTTAAAGACGGCAAGATTGACGGCATTTGCAAGGCAGGCAATAAG	314
Qy		1200	GACATGCAAGATGGCGTAAGCCCTCATATGGTGGTGGTGGGSCACAGAAGCACATAGCA	1259
Db		315	GACATGCAAGATGGCGTAGATAATAATCTTTGCGTAGGTCCTGCTACAGAGGCTTTGGCA	374
Qy		1260	GGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTCAACACCCAATTCTCTTTCTCCCA	1319
Db		375	GCTGAGGCGCTTGATTGTAAACCGCTGGTGGCATCGATACGCATATTCATTTATCTCTCCC	434
Qy		1320	CACAAATTCCTACCGCTCTAGCCCAATGGCGGTTTACACCAATGTTTGAGGCGGCACAGGT	1379
Db		435	CAA'AAATTCCTACTGCTTTTGGCAGCGGGGTTACAACCAATGATTGGAGGAGGCACAGGA	494
Qy		1380	CTGTAGATGGCACGAATGCGACTACTATCACTCCGGGCAAAATGGAACTTGCACCGCATG	1439
Db		495	CCTGCGATGGCAGATGCGACCAACCATCACTCCCGGACGCGTAATCTAAAAAGTATG	554
Qy		1440	TTGCGCGCAGAGAGATGATTTCTATGAATGTGGGCTTTTTTGGGCAAGGCAATAGCTCT	1499
Db		555	TTGCGTGCAGCCGAAGAATAACGCATGAATCTAGGCTTTTTTGGCTAAGGGGAATGTGTCT	614
Qy		1500	AGCAAAAAACA'CTGTGAGAACAGT'AGAGCGGCGCGCATTCGTTT'ATAAT'TGCATGAA	1559
Db		615	TACGAACCTCTTTTACGCGATCAGATTGAAGCAGGGCGCATTCGTTT'AAAAATCCACGAA	674
Qy		1560	GACTGGGGCAACAACCAAGTGGCATCGATCACTGTGTGACGCTGGCAGATGAATACGAT	1619
Db		675	GACTGGGGAAGCACACCTGCAGCTATTTCACCACTGCCTCAATGTGCGCGATGAATACGAT	734
Qy		1620	GTGCAAGTTTGTATCCACACCGATACAGTCAANTGAGGCGAGGTTATGTAGATGACACCGCTA	1679
Db		735	GTGCAAGTGGCTATCCACACCGATA'CCCTT'AACGAGGCGGGCTGTGTGAGAAGACACCCCTA	794
Qy		1680	AATGCAATGAACGGCGCGCATTCATGCTTACCACATTCAGGAGCGGGTGCAGAGCAC	1739
Db		795	GAGCGGATGCCGGGCGCACATCCATACCTTCCACHACTAGAGGGGCTGGGGGTGAGCAC	854
Qy		1740	TCACCTGATTTATCACCATGGCAGGCGAGCTCAATA'TTCTACCTCTCTCCACCA'CCCCC	1799
Db		855	GCTCCAGATGTTATCAAAATGGCAGGGAATTTAA'CA'TTCTACCGGCTCTCTACTAACCCG	914
Qy		1800	ACTATTTCCCTATACCATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCAC	1859
Db		915	ACCA'TTCTTTTACCACAAAA'CACT'GAAGCCGAGCA'CATGGACATGTTTA'TGGTGTGCCAC	974
Qy		1860	CACCTAGACAAA'CGCAT'CCGCGAGGATTTACAAT'TTTCT'CAAAAGCGTATCCGCCCCGGC	1919
Db		975	CAC'TTGGATAAAGTATCAAGGAAGATGTGCAGTTTGGCCGATTCGAGGATTCGCCCCCA	1034
Qy		1920	TCTATCGCGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGAT	1979
Db		1035	ACTATCGCGCTGAAGACCAACTCCATGACATGGGGATCTTTTCTAT'CACCAGCTCCGAC	1094
Qy		1980	TCCGAAGCAATGGGCGTGCAGCGGAAGTATTCCTTCGA'ACTTTGGCAGACTGCGGATTAAG	2039
Db		1095	TCTCAGGCTATGGACGCGTAGGCGAGGTGATCACACGCA'CTTGGCAGACAGCAGACAA	1154
Qy		2040	AATAAAAAAGAA'TTTCGTAA'GCTTCTCTGAAGATGGCAAGATAA'CGATAATTTCCGCA'IT	2099
Db		1155	AACAAAAAGATTTGGCGCTTGAAGAGGAAAAAGGCGATAA'CGACA'ACTTCCGCA'IT	1214
Qy		2100	AAGCGCTACATCTCCAAATA'CACTATCAACCCCGCTTTTGA'CCCA'CGCGGTGAGCGAGTAT	2159
Db		1215	AAAGGCTACATCTCTAAATACACA'CATCAACCCCGCATCGCGATTCGGGATTTCTGACTAT	1274

QY	2160	ATCGGCTCTGTGGAAAGAGGCGAAGATCGCCGACTTGGTGGTGTGGAAATCTGCCTTTT	2219
Db	1275	GTGGGCTCTGTGGAAAGTGGGCAATATACGCCGACCTCGTGTCTTTGGAGTCCGGCTTTCCTT	1334
QY	2220	GGCGTAAACCCAAAATCGTGATCAAAGGCGGTATGGTGGCTCTCTCTGAAATGGGCGAT	2279
Db	1335	GGCATTAAGCCCAATATGATTTATTAGGGCGGATTTATTGGCTCTCTCAATGGGCGAT	1394
QY	2280	TCTAACCGGCTGTGTCACATCTCCCAACCGGTTTTATTACCGGAAATGTTTTGGGCATCAC	2339
Db	1395	GCCAATCGCTATTTCGCACCCCTCAGCCGCTGTATTACCGTGAAATGTTTGGACACCAT	1454
QY	2340	GGCAAGCGGAAATTTTGACACAGCATCACATTTTGTTCCAAAGTCGCTATGAAAATGGC	2399
Db	1455	GGGAAAACAAATTCGACACAATATCACATTTTGTGTCCAGGGGTTACAGGCAAGG	1514
QY	2400	GTGAAGAAAGCTGGGCTTTAGAGCGCCAAAGTTCTACCGGTCAAAAATCGCCGTAAACATC	2459
Db	1515	ATCAAGAAGAACTTAGGGCTAGATCGGCTGGTATTGCCAGTGAAAAACCTGTCGAATATC	1574
QY	2460	ACCAAGAAAGACTTCAAGTTCAAGHCAAAACGGCAAAAATCACCGTCGATCCGAAACC	2519
Db	1575	ACTAAAAGGACCTCAAATTCACAGATGTACCGCAATATTGATGTCAACCCCTGAAACC	1634
QY	2520	TTCCAGGCTTTTGTAGATGTCAAACTCTGCACCTCTAAAACCCACTCGCAAGTGCCTCTA	2579
Db	1635	TATTAAGGTGAAGTGGATGCAAGAGGATTAACCTCTTAAGCAGCAGATGAATTGAGCCTA	1694
QY	2580	GCCAGCGCTACACTTTCTTCTAG	2603
Db	1695	GCGCACTTTATAATTTGTTCTAG	1718

RESULT 9

US-10-335-977-3849
; Sequence 3849, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 3849:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...1710
SEQUENCE DESCRIPTION: SEQ ID NO: 3849:
US-10-335-977-3849

Query Match 30.8%; Score 886.4; DB 8; Length 1710;
Best Local Similarity 70.0%; Pred. No. 2.9e-184;
Matches 1193; Conservative 0; Mismatches 511; Indels 0; Gaps 0;

QY 900 AAAATGAAAAAACAAGATATGTAATACCTTACCGACCCACCAAGGCGATAAAGTGGC 959
DB 7 AAGATTAGCAGAAAGAATATGTTCTATGTATGGCCCTACACAGGCGATAAAGTGAGA 66

QY 960 TTAGGAGATACCGATCTTTGGGCGAAGAGTAAACATGACTATACACCTATGGGCGAAGAA 1019
DB 67 TTGGCGATACAGATTGATCGCTGAAGTAAACATGACTATACACCTATTTATGGCGAAGAG 126

QY 1020 CTTAAATTTGGCGGGTAAACATATCGTGAGGGTATGGGTGAGCAATAGCCCTGAT 1079
DB 127 CTTAAATTTGGCGGGTAAACATATCGTGAGGGTATGGGTGAGCAATAGCCCTGAT 186

QY 1080 GAAACACCTAGATTAGTCACTCACTAAACCGGATGATTATCGACTACACCGGATTTAC 1139
DB 187 AAAGAAGAACTGGATTGATCATCACTAAACCGGATGATTATCGACTACACCGGATTTAT 246

QY 1140 AAAGCCGACATTTGGATTAAACCGGAAATCATGGCATTTGGCAAGGCGAAGAAACAG 1139
DB 247 AAAGCGATATTGTTATTAAGACGCGAAATCGTGGCATTTGTAAGGCGGTAAACAA 306

QY 1200 GACATGCAAGATGCGTAAGCCCTCATATGTCGTGGGTGTTGGGCGAAGAGCACTAGCA 1259
DB 307 GACATGCAAGATGCGTTAAACCAATCTTAGCGTGGGTCTGCTACTGAAAGCGCTAGCT 366

QY 1260 GGGGAAGTATGATTATACCGCTGGGGGATTCATTCACACCGACTTCTTCTTCCA 1319
DB 367 GGTGAAGGCTTAATCGTAAGCGCTGGTGTATGACACACATCCACTTCATTTCAACC 426

QY 1320 CAACAAATTCCTACCGCTTAGCCCAATGGCGTTACACCATGTTTGGAGGGCGCACAGGT 1379
DB 427 CAACAAATTCCTACAGCTTTTGAAGCGGTGTAACACCATGTTTGGTGGCGAACCGGT 486

QY 1380 CTTGTAGATGCGAAGTACGACTACTATCACTCCGGGCAAAATGGAATTCACCGCATG 1439
DB 487 CTTGTAGTGCACATAATGCGACTACTATCACTCCAGGTAGAGAAACTTAAATGGATG 546

QY 1440 TTGGCGCGAGAGAGTATCTTATGAATGTGGCTTTTGGGCAAAAGGCAATAGCTCT 1499
DB 547 CTCAGAGCGGCTGAAGAATATCTTATGAATCTTAGGTTCTTAGCTTAAAGGTAAAGCTTCT 606

QY 1500 AGCAAAAAAACAATTTGTAACAGATGAGCGGGCGGATGTTTAAATTCGATCAA 1559
DB 607 AACGACGCGAGCTTAGCGGATCAAAATTAAGCGCGGTGCGATTGAGTTTAAATTCACGAA 666

QY 1560 GACTGGGGCGCAACACCAAGTGCATGATCACTGCTTGGCGTGGCGAGATGAATACGAT 1619
DB 567 GACTGGGGCGCAACATCTCTTGGCAATCAATCATGCTTGGAGTGGCGACAAATACGAT 726

QY 1620 GTCAAGTTTGTATCCACACCGATCACTCAATGAGGCGAGTTATGTAAGTACACCCCTA 1679
DB 727 GTCAAGTTGCTATCCACACAGACACTTTGAATGAAGCGGTGCGTGGTGAAGACACTATG 786

QY 1680 AATCAATGAACGGCGGCGCATCCATCCCTACCACTTGGAGGCGGGTGGAGACAC 1739
DB 787 GCAGCTATTGCGGACGCGACTATGACACTTTTCCACTGAAGGCGCTGGTGGCGGACAC 846

QY 1740 TCACCTGATGTTATCACCATGGCGGCGAGCTCAATATTTCTACCTCCTCCACCAACCCCC 1799

DB 847 GCTCTGACATTTATTAAGTGGCGGTGAACAACATCTTACCCGCTTCCATAACCCC 906

QY 1800 ACTATTCCTATATACCATTAATACGCTTGAGAACACTTTAGACATGCTCATGACATGCCAC 1859

DB 907 ACTATCCCTTCTCACTGTGATACAGAGCCGAACATGACATGCTCATGGTGTGCCAC 966

QY 1860 CACCTAGACAAACGATCCGGAGAGATTAAATTTTCTCAAGCCGATTCGCGCCCGGG 1919

DB 967 CACTTGGATAAAAGCATTTAAAGAGATGTCAGTTCGCTGATTAAGGATCCGCGCTCAA 1026

QY 1920 TCTATCGCGCTGAGATGCTCTCATGATATGGGTGATCGCGATGACAGCTCGGAT 1979

DB 1027 ACCATTCGCGCTGAAGACACTTTGATGACATGGGATTTTCTCAATCAGGTTCTGAC 1086

QY 1980 TCGAAGCAATGGGCGGTGCGAGCGAAGTATCTCGAACTTGGCAGACTGGCGATAAG 2039

DB 1087 TCTCAAGCGATGGGTGCTGTGGTGAAGTTATCACTAGACTTGGCAACAGCTGACAAA 1146

QY 2040 AATAAAAAAAGATTTGGTAAGCTTCTGAAGATGGCAAGATAACGATATTTCCGAT 2099

DB 1147 AACAAAAAAGATTTGGCGCTTGAAGAGAAAGAGCGATAACGACAACTTCAGGATC 1206

QY 2100 AAGCGTACATCTCAAAATACACTATCAACCCGCTTTGACCCAGCGCTGACGAGTAT 2159

DB 1207 AAACGCTACTTGTCTAAATATACACCATTAACCCAGCGATCGCTCATGGGATTAGCGAGTAT 1266

QY 2160 ATCGCTCTGTGGAAGAGGGCAAGATGCGCGACTTGGTGGTGGGAATCTCTGCTTTT 2219

DB 1267 GTAGGTTCTGTAGAAAGTGGGCAAGTGGCTGACTTGGTATTTGGAGTCCAGATTTCT 1326

QY 2220 GGCGTAAACCCAAATCGTGATCAAGCGGTATGGTGGTCTTCTCTGAAATGGGCGAT 2279

DB 1327 GGCGTAAACCCAAATCATGATCAAGAGCGGATTCATTGCGTTAAGCCAAATGGGCGAT 1386

QY 2280 TCTAAGCGCTGTGCGCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCGATCAC 2339

DB 1387 GCGAAGCGCTTCTATCCCTACCCCAACCGGTTTATACAGAGAAATGTTTCGCTCACCAT 1446

QY 2340 GCGAAGCGGAAATTTGACACGAGCATCACTTTTGTTCGAAAGTCCGCTATGAAATGGC 2399

DB 1447 GGTAAAGCTAAATAGATGCAAAACATCACTTTTGTCTCAAGCGCTTATGCAAGGC 1506

QY 2400 GTGAAGAAAAAGCTCGGCTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCCGTAAATC 2459

DB 1507 ATTAAGAGAAATTTAGGCTTGAAGACAAAGTGTGCGCGTTAAATAATTCGAAACATC 1566

QY 2460 ACCAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAAATCACCGTCGATCCGAAACC 2519

DB 1567 ACTAAAAAAGACATGCAATTTCAATGACACTACCGCTCAATTTGAAGTCAATCTCGAACT 1626

QY 2520 TTCGAGCTTTGTAGATGGCAAACTCTGCACTCTTAAACCCACCTCGGAAGTGCCTCTA 2579

DB 1627 TACCATGTTGCTGGATGGCAAGAGTCACTTCTAAACCCAGCTAATAAAGTGAGCTTG 1686

QY 2580 GCCAGCGCTACACTTCTTCTAG 2603

DB 1687 GCTCACTCTTTAGCATTTTCTAG 1710

RESULT 10
US-09-895-913A-251
; Sequence 251: Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1 Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome

FILE REFERENCE: 06132/043002
CURRENT APPLICATION NUMBER: US/09/895,913A
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 08/881,227
PRIORITY FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 251
LENGTH: 1815
TYPE: DNA
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: CDS
LOCATION: (51)...(1757)
US-09-895-913A-251

Query Match 30.8%; Score 885.8; DB 3; Length 1815;
Best Local Similarity 70.0%; Pred. No. 4.1e-184;
Matches 1193; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

QY 900 AAATGAAAAAACAAGAAATATGTAATACCTACCGAACCCACAAAGGCGATAAAGTGC 959
DB 57 AAGATTAGCGAAGAAATATGTTCTATGATGCGCCCTACTACAGGCGATAAAGTGAGA 116
QY 960 TTAGGAGATACCGATCTTTGGCGAAGATGAAACATGACTATACCATCTATGGCGAAGAA 1019
DB 117 TTGGGCGATACAGACTTGTATCGCTGAAGTAGAATGATGATGATGATGATGATGATGAT 176
QY 1020 CTTAAATTTGGCGGGGTAAAACTATCCGTGAGGATGATGGGTGAGAGCAATAGCCCTGAT 1079
DB 177 CTTAAATTCGGTGGCGGTAAACCTGAGAGAGGCGATGAGCCATCCNACACCTTAGC 236
QY 1080 GAAAAACCCCTAGATTAGTATCATCACTAACCGCGATGATTATCGACTACACCGGATTTAC 1139
DB 237 AAAGAAGAAATTTGATCTAATCATCACTAACCGCTTTAATCGTGGATTAACCGGTATTAT 296
QY 1140 AAAGCCGACATTTGGGATTAATAAAGCGGAAATCATATGGCATTTGGCAAGGCGAAGAAACAG 1199
DB 297 AAAGCGGATATTGGTATTAAGATGGAATAATCGCTGGCATTTGTTAAAGGCGGTAAACAA 356
QY 1200 GACATCAAGATGCGGTAAAGCCCTCATATGTCGTGGGTGTTGGCGACAGAGCACTAGCA 1259
DB 357 GACATCAAGATGCGGTAAAGCAAACTTTAGCGTAGGTCCTGCTACTGAAGCCCTTAGCC 416
QY 1260 GGGGAGGATGATTATTAACCGCTGGGGGAATCGATTACACACCCCACTCTCTTTCTCCA 1319
DB 417 GGTGAAGTTTGTATGCTAACTGCTGGGTGATTGACACACATCCACTTCATTTACCC 476
QY 1320 CAACAAATTCCTACCGCTTAGCCAAATGGCGTTTCAACACCATGTTTGGAGGCGGCACAGGT 1379
DB 477 CAACAAATTCCTACAGCTTTTGAAGCGGTGTAACCAACCATGATTGTTGGCGGAACCTGGT 536
QY 1380 CTTGTAGTGGCAGATGCGACTATCATCTACCTCGGGGCAATGGAACTTGCACCGCATG 1439
DB 537 CTTGTGTATGGCACTAATGCGACTACTATCACTCCAGCGAGAGAAATTTAAATGGATG 596
QY 1440 TTGCGCGCAGCAGAGATTTCTATGAATGTGGGCTTTTGGCAAGGCAATAGCTCT 1499
DB 597 CTGAGAGCGGCTGAAGAAATTTCTATGAATTTAGTTTCTTGGCTAAGGTTAAGCTTCT 656
QY 1500 AGCAAAAAACAATTTGAGAAACAAGTAGAGCGGCGCGATTTGGTTTTAAATTCATGAA 1559
DB 657 AACGACGCGAGCTTAGCCGATCAATTTGAAGCTGGTGGATTTGGCTTTAAATCCAGAA 716
QY 1560 GACTGGGCGCAACACCAAGTGGATGATCACTGTTGAGCGGTGGCGATGATGATGAT 1619
DB 717 GACTGGGCGCAACACTCTCTCTGCAATCAATCATCGTTAGATTGTTGACAGCAAAATAGCAT 776
QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGTTATGTAGATGACACCTA 1679
DB 777 GTGCAAGTTCGTATCCACACAGACATTTGAAATGAAGCCGGTTGCGTGAAGACATATG 836
QY 1680 AATGCAATGAACGGGCGCGCCATCCATCCCTACCATTTGAGGAGCGGTGGAGGACAC 1739

RESULT 11

US-10-282-122A-22427
; Sequence 22427, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert

DB 837 GCAGCTATTGCCGAGCGCACTATGCACATTTCCACACTGAAGGTGCTGCGGGGACAC 896
QY 1740 TCACCTGATGTTATCACCATGCGAGCGAGCTCAATATTCTACCCCTCTCCACACACCCCC 1799
DB 897 GCTCCTGATATTATTAAAGTAGCTGGTGAACAACAATTTCTCCCGCTTCCACTAACCCC 956
QY 1800 ACTATTCCCTATACCATTAAATAGGTTGCAGAACACTTAGACATGCTCATGACATGCGAC 1859
DB 957 ACTATCCCTTTTCACTGTGAATACAGAAAGCAGAACACATGAGACATGCTTATGGTGTGCG 1016
QY 1860 CACCTAGACAAAACGATCCGCGAGGATTTACAAATTTTCTCAAAAGCCGTATCCGCGCCCGC 1919
DB 1017 CACTTGGATAAAGCATTTAAAGAGATGTTCAAGTTCGCTGATTTCAAGGATCCGCGCTCAA 1076
QY 1920 TCTATCGCGCTGAAGATGCTCTCCATGATATGGGTGATCGCGATGACAAAGCTCGGAT 1979
DB 1077 ACCATTGGCGCTGAAGACACTTTTGCATGACATGGGGATTTTCTCAATCACCAGCTCTGAC 1136
QY 1980 TCGAAGCATATGGGCGTGCAGGCGAAGTATCTTCGAACTTTGGGAGAGCTGCGGATAAG 2039
DB 1137 TCTCAAGCTATGGGTGCTGTTGGGTGAAAGTTATCACTAGAACTTTGGCAAAACAGCTGACAAA 1196
QY 2040 AATAAAAAGAAATTTGGTAAAGCTTCTGAGATGGCAAGATTAACGATATTTCCGCAAT 2099
DB 1197 AACAAAAGAAATTTGGCGCGCTTGAAGAAGAAAGGCGGATAACCACTTCAGGATC 1256
QY 2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGCGTGAGCGAGTAT 2159
DB 1257 AAACGCTACTGCTTAATACACCATTAACCCAGCGATCGCTCATGGGATTAAGGAGTAT 1316
QY 2160 ATCGCTCTGTGGAAGAGGCGAAGATCGCCGACTTTGGTGGTGGAAATCTGCTCTTTT 2219
DB 1317 GTAGTTCTGTAGAAAGTGGGCAAGTGGCTGACTTTGGTATTTGGAGTCCCGCAATCTTT 1376
QY 2220 GCGGTAAACCCCAATCGTATCAAGCGCGTATGGTGTCTTCTCTGAAATGGGCGAT 2279
DB 1377 GCGGTAAACCCCAATCATGATCATCAAGCGCGGTTCAITGCGTTGAGTCAAAATGGGCGTAC 1436
QY 2280 TCTAACGCGTCTGTGCCCACTCCCAACCGGTTTATACCGCGAAATGTTTGGGCGATCAC 2339
DB 1437 GCGAAGCTTCTATCCCTACCCCAACACAGTTTATACAGAGAAATGTTTCGCTCATCAT 1496
QY 2340 GCGAAGCGAAATTTGACACACGACATCACTTTTGTTCCTCAAAAGTCCGCTATGAAATGGC 2399
DB 1497 GGTAAAGCCAAATACGATGCAACATCACTTTTGTCTCAAGCGGCTTATGACAAAGGC 1556
QY 2400 GTGAAAGAAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAACCTGCGCGTAAACATC 2459
DB 1557 ATTAAGAAAGAAATTAGGGCTTGAAGACAAAGTGTTCGCGGTAAAAAATTTGCAGAAACATC 1616
QY 2460 ACCAAGAAAGACTTCAAGTTCAACGACAAAGCGCAAAATCACCCTCGATCCGAAACCC 2519
DB 1617 ACTAAAAAGACATGCAATTTCAACGACACTACCGCTCAATTTGAAGTCAATCTGAAACT 1676
QY 2520 TTCGAGGTCTTTGATAGTGGCAAACTCTGCACTCTTAAACCCACCTCGCAAGTGCCTCTA 2579
DB 1677 TACCATGTTGTTGGATGGCAAGAGTAACTTCTAAACGAGCAATTAAGATGAGCTTG 1736
QY 2580 GCCGAGCGCTACACTTTCTCTTAGG 2604
DB 1737 GCGCAACTCTTTAGCATTTCTTAGG 1761

APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 22427
 LENGTH: 1710
 TYPE: DNA
 ORGANISM: Helicobacter pylori
 US-10-282-122A-22427

Query Match 30.7%; Score 884.8; DB 8; Length 1710;
 Best Local Similarity 70.0%; Pred. No. 6.5e-184;
 Matches 1192; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

Qy	900	AAAATGAAAACAAAGATATGTAATACCTACCGAACCCCAAGCGGATAAAGTGGC	959
Db	7	AAGATTAGCAAGAAATATGTTCTATGTATGGCCCTACTACAGCGGATAAAGTGGA	66
Qy	960	TTAGGAGATACCGATCTTTGGGCAGAACTAGAAATGACTATACCATCTATGGCGAAGAA	1019
Db	67	TTGGCGGATACAGACTTGATCGCTGAAGTAGAATGACTACCATTTTATGGCGAAG	126
Qy	1020	CTTAAATTTGGCGGGTAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT	1079
Db	127	CTTAAATTCGGTGGCGTTAAACCTGAGAGAGGATGAGCAATCAACACCCCTAGC	186
Qy	1080	GAACAACCCCTAGATTTAGTCTATCACTAAACCGGATGATTCGACTACACCGGATTTAC	1139
Db	187	AAAGAAGAAATTTGGATCTAATCATCACTAACCGCTTTAATCGTGGATTAACCGGTTAT	246
Qy	1140	AAAGCCGATTTGGATTTAAAGAGGCAAAATCCATGGCATTTGGCAAGGCAAGAAACAG	1199
Db	247	AAAGCGGATTTGGTATTAAGATAGGCAAAATCGCTGGCATTTGGTAAGCGGTAACAA	306
Qy	1200	GACATGAAGATGGCGTAAGCCCTCATATGCTCGTGGGTGGGCGACAGAACTAGCA	1259
Db	307	GACATGAAGATGGGTTAAAGCAATCTTAGCGTAGTCTGCTACTGAAGCCCTTAGCC	366
Qy	1260	GGGAAGGTATGATTTATCCGCTGGGGAAATCGATTACACACCCACTTCTCTTCTCA	1319
Db	367	GGTGAAGGTTTATCGTAACTGCTGGTGTATTTGACACACATCCACTTTCATTTCA	426

Qy	1320	CAACAAATCCCTACCGCTCTAGCCAAATGGCGTTACAAACCATGTTTGGAGCGCGCAGGT	1379
Db	427	CAACAAATCCCTACAGCTTTTTCAGAGCGGTGTAAACCAACCATGATTTGGTGGCGAATGGT	486
Qy	1380	CCTGTAGATGGCAAGATGCGACTACTACTCCGGGCAAAATGGAACCTTTGACCGCATG	1439
Db	487	CCTGTAGTGGCACTAATGCGACTACTACTCCAGGCAAGAAATTTAAATGATG	546
Qy	1440	TTGCGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAGCAATAGTCTCT	1499
Db	547	CTCAGAGCGGCTGAAGAAATTTCTATGAATTTAGGTTTCTTGGCTAAAGTAAGCTTCT	606
Qy	1500	AGCAAAAACAACTTGTAGAACAGTAGAGCGGGCGGCAATGCTTTTAAATGCAAGAA	1559
Db	607	AACGACGCGAGCTTAGCCGATCAAAATTTGAAGCTGGTGGCTTTTAAATCCAGAA	666
Qy	1560	GACTGGGGCACAAACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATAGAT	1619
Db	667	GACTGGGGCACCACTCTCTTCTGAATCAATCATGCTTAGATGTTTGCAGCAATAGAT	726
Qy	1620	GTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGGCAAGTTTATGTAGATGACCCCTA	1679
Db	727	GTGCAAGTGGCTATCCACACAGACACTTTGAATGAAGCGGTTTGGCTGAAGACACTATG	786
Qy	1680	AATGCAATGAACGGCGGCGCATCCATGCTTACCACTTGAAGGAGCGGTGGAGACAC	1739
Db	787	GCAGCTATTTGGCGGACGCACTATGCACACTTTCCACACTGAAGGTGCTGGCGGAGCAC	846
Qy	1740	TCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCCACACCCCC	1799
Db	847	GCTCTGATATTTAATTAAGTAGCTGTGAACACAACTTTCTCCCGCTTCCACTAACCCC	906
Qy	1800	ACTATTTCCCTATACCATTAAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCAC	1859
Db	907	ACTATTTCCCTTCACTGTGATACAGAACAGAACATGACATGCTTATGTTGTGCCAC	966
Qy	1860	CACCTAGACAAACGATCGCGAGGATTTCAATTTCTCAAAGCCGTATCCGCCCGGC	1919
Db	967	CACCTTGGATAAAGCATTAAGAGAGATGTTCAAGTTCGCTGATTAAGGATCCGCCCTCAA	1026
Qy	1920	TCTATCGCGGCTGAAGATGCTCCATGATATGGTGTGATCCGATGACAAAGCTCGAT	1979
Db	1027	ACCATTTGGCGCTGAAGACACTTTTGCATGACATGGGGAATTTCTCAATCACCGCTGTAC	1086
Qy	1980	TGCAAGCAATGGCGGTGCGAGCGAAGTATTCCTCGAACTTTGGCAGACTCGCGGTAAG	2039
Db	1087	TCTCAAGCTATGGTGTGGTGGTGAATTTCACTAGAACTTTGGCAACAGCTGACAA	1146
Qy	2040	AATAAAAAGAAATTTGGTAAGCTTCTGAAAGATGGCAAGATAACGATATTTCCGCAT	2099
Db	1147	AACAAAAAGAAATTTGGCGCTTGAAGAGAAAGGCGGATAACGACAACTTCAGGATC	1206
Qy	2100	AAGCGGTACATCTCAAAATACATATCAACCCGCTTTGACCCACCGCTGAGCGGATAT	2159
Db	1207	AAACGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAGCGATAT	1266
Qy	2160	ATCGGCTGTGGAAGAGGCAAGATCGCGGATTCGCTGGTGTGGTGAATCTCGCTTTT	2219
Db	1267	GTAGGTTCTGTAGAAGTGGGCAAGTGGCTGACTTGTGATTTGTGGAGTCCCGCATTTCT	1326
Qy	2220	GGCGTAAACCCCAAAATTCGTGATCAAGGCGGTATGGTGGTCTTCTCTGAAATGGGCGAT	2279
Db	1327	GGCGTAAACCCCAACATGATCATCAAGGCGGTTTCATTTGGTGTAGTCAATGGGTGAC	1386
Qy	2280	TCTAAGCGCTGTGGCCACTCCCAAACCGGTTTATACCGCGAAATGTTTGGGCGATCAC	2339
Db	1387	GGCAACGCTTCTATCCCTACCCCAACCAAGTTTATTACAGAGAAATGTTTCGCTCATCAT	1446
Qy	2340	GGCAAGGCAAAATTTGACACAGCATCACTTTTGTTCCAAAGTCCGCTATGAAATGGC	2399
Db	1447	GGTAAGGCCAAATACGATGCAAAACATCACTTTTGTGTCTCAAGCGGCTTATGACAAAGC	1506
Qy	2400	GTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAACTGCCGTAAATC	2459

Db 1507 ATTAAGAGAGATTAGGCTTGAAGACAGAGTTGCGCGTAAATAATGCGAAGAACATC 1566
Qy 2460 ACCAAGAAAGACTTCAAGTTCAACGACAAAACGCGCAAAAATACCGTGCATCCGAAACC 2519
Db 1567 ACTAAAAAAGACATGCAATTCACGACACTACCGCTCACATTTGAAGTCAATCCTGAAACT 1626
Qy 2520 TTCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACTCCGAGTGCCTCTA 2579
Db 1627 TACCATGTGTTGCGGATGGCAAGAAGTAACCTTCTAAACGAGCAATATAAGTGAGCTTG 1686
Qy 2580 GCCCAGCGCTACACTTTCTCTAG 2603
Db 1687 GCGCACTCTTAGCATTTTCTAG 1710

RESULT 12

US-10-500-447A-5
; Sequence 5, Application US/10500447A
; Publication No. US20050150016A1
; GENERAL INFORMATION:
; APPLICANT: PARK, Hee-Sung
; TITLE OF INVENTION: Method for producing a recombinant protein using pollen
; FILE REFERENCE: YLOP040518US/PCT
; CURRENT APPLICATION NUMBER: US/10/500,447A
; CURRENT FILING DATE: 2004-06-30
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-500-447A-5

Query Match 30.5%; Score 880; DB 10; Length 1710;
Best Local Similarity 69.8%; Pred. No. 7.5e-183;
Matches 1189; Conservative 0; Mismatches 515; Indels 0; Gaps 0;

Qy 900 ABAATGAAAACAGAAATATGTAATACCTACGACCCCAACCAAGCGGATAAAGTGGC 959
Db 7 AAGATTAGCAGAAAGAAATATGTTCTATGATGGCCCTACTACAGCGGATAAAGTGAGA 66
Qy 960 TTAGGAGATACCGATCTTTGGGAGAGTGAACATGACTATACCATACCTATGGCGAAGAA 1019
Db 67 TTGGCGATACAGACTTGTATGCGTGAAGTAGAATGATACCATTTATGGCGAAGAG 126
Qy 1020 CTTAAATTTGGCGGGTAAACCTATCCGTAGGGTATGGGTACAGACAAATAGCCCTGAT 1079
Db 127 CTTAAATTTGGCGGGTAAACCCCTAAGAGAAGGCATGAGCCAATCTAACACCCCTAGC 186
Qy 1080 GAAAAACCCCTAGATTAGTATCACTAACCGGATGATTATGACTACACCGGATTTAC 1139
Db 187 AAAGAAGAACTGGATCTCAATCATCACTAACCGCTTTATTCGTGGATTACACCGGTATTAT 246
Qy 1140 AAAGCGACATTTGGATTAAAAACGCAAAATCCATGCGCATTTGCAAGGCGAAGAACAG 1199
Db 247 AAAGCGGATTTGTTATTAAGATGGCAAAATCGTGGCAATTTGTAAGGCGGTAAACAA 306
Qy 1200 GACATCAAGATGCGGTAAAGCCCTCATATGGTGGGTGTGGGCAAGAGCACTAGCA 1259
Db 307 GACATCAAGATGCGGTAAAGCAATCTTAGCGTGGTCTCTACTTAGAGCCTTAGCC 366
Qy 1260 GGGGAAGGTATGATTTATCCGCTGGGGGAATCGATTCAACACCCACTTCTTTCTCCA 1319
Db 367 GGTGAAGGTTTGTATCGTTAACTGCTGGTGGTATTGACACACATCCACTTCACTCTCCCC 426
Qy 1320 CAACAATTCCTACCGCTCTAGCCATGGGTTTACACCATGTTTGGAGGGCGCACAGGT 1379
Db 427 CAACAATTCCTACAGCTTTTGAAGCGGTGTAACACAGATGATTGGTGGCGAAGTGGC 486
Qy 1380 CTTGTAGATGGCAAGTACGACTACTATCACTCCGGGCAATGGAATCTTGCCACCGCATG 1439

Db 487 CTGCTGATGGCACTTAACGAAACCCTATCACTCCAGGTAGAGAAATTTAAATGTGATG 546
Qy 1440 TTGCGCGCAGCAGAGAGATTCTATGAATGTGGGCTTTTGGGCAAAAGCAATAGCTCT 1499
Db 547 CTCAGAGCGGCAGAGAAATTTCTATGAATCTTAAGTTTCTAGCTAAAGGTAAAGCTTCT 606
Qy 1500 AGCAAAAACAACTTTGTAGAACAGAGTGAAGGGGGCGCATTTGGTTTAAATTTGCATGAA 1559
Db 607 AACGATGCAAGCTTAGCCGATCAAAATTTGAAGCGGTGCGATTTGGCTTTTAAAAATCCAGAA 666
Qy 1560 GACTGGGGCACAACCAACCAAGTGGATCACTGCTTTCAGGGTGGCAGATGAATAGAT 1619
Db 667 GACTGGGGCACCACCTCTTCTGCAATCAATCATGCGTTAGATGTTGCGGACAAATAGAT 726
Qy 1620 GTGCAAGTTTGTATCCACACCCGATACAGTCAATAGGCGCAGGTTATGTAGATGACACCCCTA 1679
Db 727 GTGCAAGTCGCTATCCACACACACTTTGAATGAAGCCGTTGTGTAGAGACACTATG 786
Qy 1680 AATGCAATGAAGCGGCGCCCATCCATGCTTACCACTTGAAGGAGCGGTGGAGACAC 1739
Db 787 GCAGCCATTTGCGGACGCACTATGCACTTTCCACACTGAAGCGCTGGTGGCGACAC 846
Qy 1740 TCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACCACCCCT 1799
Db 847 GCTCCTGATATTTAATAAGTAGCTGGTGAACACACATTTCTGCCGCTTCCATAACCCCT 906
Qy 1800 ACTATTCCTATACCAATTAATACGTTGCGAGAACACTTAGACATGCTCATGACATGCCAC 1859
Db 907 ACTATTCCTTTCTACTGTGATACAGAGCAAGAACATGACATGCTTATGGTGTGCCAC 966
Qy 1860 CACTTAGACAAACGCAATCCGCGAGGATTTACAATTTTCTCAAGCGGTATCCGCCCGCGC 1919
Db 967 CACTTGGATAAAGCAATTAAGAGAGATGTTCAAGTTTCGCTGATTTCAAGGATCCGCCCTCA 1026
Qy 1920 TCTATCGCGCTGAAGATGCTCTCATGATATGGTGTGATCGCATGACATGACAGCTCGAT 1979
Db 1027 ACTATTTGCGGCTGAAGACACTTTGATGACATGGGGATTTTCTCAATCACCAGTTCTGAC 1086
Qy 1980 TCGAAGCAATTTGGGCGTGGCGGCGAAGTATCTCGAACTTTGCGACAGCTCGGATAAG 2039
Db 1087 TCTCAAGCTATTTGGGTCGTGGTGAAGTATCACAGAACTTTGGCAACAGCTGACAA 1146
Qy 2040 AATAAAGAAATTTGGTAAGCTTCTCAAGATGGCAAGATAACGATAATTTCCGCAT 2099
Db 1147 AACAAAGAAATTTGGCGCTTGAAGAGAAAGGCGATAACGACACTTTCAGGATC 1206
Qy 2100 AAGCGCTATCTTCAAAATACATATCAACCCCGCTTTGACCCACGCGGTGAGCGGATAT 2159
Db 1207 AAAGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAAGCGAGTAT 1266
Qy 2160 ATCGCTCTGTGGAAGAGGCGAAGATCCGCACTTGGTGGTGGTAATCTTCCCTTTT 2219
Db 1267 GTAGGTTCTGTAGAGTGGCGAAGTGGCTGACTTGGTGTGGTGGTGGTGGTGGTGGT 1326
Qy 2220 GGCGTAAACCCCAAAATCGTATCAAGCGCGTATGGTGGTCTTCTCTGAAATGGCGCAT 2279
Db 1327 GGCGTGAACCCCAACATGATCATCAAGCGCGATTCATTTGATTTGATTTGATTTGATTT 1386
Qy 2280 TCTAACCGCTGTGTGCCACTCCCCAACCGGTTTATTAACCGCAAAATGTTTGGGCAATC 2339
Db 1387 GCGAAGCTTCTATCCCTTACCCCAACCGGTTTATTAAGAGAAATGTTTCGCTCATCAT 1446
Qy 2340 GGCAAGCGAAATTTGACACGAGCATCTTTTGTTCGTTTCAAGTGGCGCTATGAAATGCG 2399
Db 1447 GGTAAAGCTAAATACGATGCAAAATCACTTTTGTCTCAAGCGGCTTATGACAAAGCG 1506
Qy 2400 GTGAAGAAAGCTGGGCTTAGAGCGCGCAAGTCTTACCGGTCAAAAACCTGCGGTAAACATC 2459
Db 1507 ATTAAGAGAAATTAGGCTTTGAAGGCAAGTGTTCGCGGTAAATAATTTGCAAAACATC 1566
Qy 2460 ACCAAGAAAGACTTCAAGTTTCAACGACAAACGCGCAAAATCACCGTGCATCCGAAAAAC 2519

Db 1567 ACTAAAAAGACATGCAATTCAACGACACTACCGCTCACATTGAAGTCAATCCTGAAACT 1626
Qy 2520 TTGAGGTCTTTGATGATGGAAGAACTGCGACCTCTAAACCCACCTCCAGTGCCTCTA 2579
Db 1627 TACCATGTGTCGTGGATGCGAAGAGTAATCTTAAACGAGCAATAAAGTGAGCTTG 1686
Qy 2580 GCCAGCGCTACACTTTCTCTTAG 2603
Db 1687 GCACAACTTTAGCATTTCTTAG 1710

RESULT 13
US-10-476-313-11
; Sequence 11, Application US/10476313
; Publication No. US20040241175A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, BRENT
; APPLICANT: BOYLE, JEFFEREY
; APPLICANT: LEW, ANDREW
; TITLE OF INVENTION: Antigen Targeting
; FILE REFERENCE: BWP-002
; CURRENT APPLICATION NUMBER: US/10/476,313
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PRS241
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 11
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-476-313-11

Query Match 30.4%; Score 876.8; DB 9; Length 1717;
Best Local Similarity 69.7%; Pred. No. 3.8e-182;
Matches 1187; Conservative 0; Mismatches 517; Indels 0; Gaps 0;

Qy 900 AAAATGAAAAAAGAAATATGTAATACCTACCGGACCCACCAAGGCGATAAAGTGCGC 959
Db 13 AAGATTAGCAGAAAGAAATATGTTCTATGATGCGCCCTACTACAGGTGATAAAGTGAGA 72

Qy 960 TTAGAGATACCGATCTTTGGGCGAGAGTAGAAGATGATACATGATACACCTATGGCGAAGAA 1019
Db 73 TTGGCGGATACAGACTTTGATCGCTGAAGTAGAAGATGATACATGATACACCTATTTATGGCGAAGAG 132

Qy 1020 CTTAAATTTGCGCGGGTAAACTATCGTCAGGATGAGGTGAGCAAGCAATAGCCCTGAT 1079
Db 133 CTTAAATTTGCGCGGGTAAACCCTAAGAGAGGCGATGAGCCAACTTAAACACCCCTAGC 192

Qy 1080 GAAACACCCCTAGATTTAGTCACTCACTAACGCGATGATTATCGACTACACCGGATTTAC 1139
Db 193 AAGAGNACTGGATCTAATCACTCACTAACCGTTTAACTCGTGATTAACCGGTATTTAT 252

Qy 1140 AAAGCGCATTTGGATTTAAAGCGGCAAAATTCATGCGATTTGGCAAGCGAGGAAACAG 1199
Db 253 AAAGCGGATATGGTATTTAAAGATGGCAAAATCGTGGCATTTGGTAAAGCGGTAACAAA 312

Qy 1200 GACATGCAAGATGGGTAAAGCCCTCATATGTCGTGGGTGGGTCAGACAGAGCACTAGCA 1259
Db 313 GACATGCAAGATGGGTAAAGCAAACTTTAGCGTGGGTCTGCTACTGAAAGCCCTAGCC 372

Qy 1260 GGGGAAGGTATGATTTATACCGCTGGGGGAATCGAATTCACACACCCACTTCTCTTCTCCA 1319
Db 373 GGTGAAGTTTGAATCGTAACTGCTGGTGGTATGACACACATCCACTTCAATTTACCC 432

Qy 1320 CAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTTAAACCATGTTTGGAGGCGGCAAGT 1379
Db 433 CAACAATTTCCCTACAGCTTTTGAAGCGGTGTAAACCAATGATTTGGTGGCGGAAGTGGT 492

Qy 1380 CTTGTAGATGCCAATTCGCACTACTACTACTCTCGGGCAAAATGGAATTTGACCGCATG 1439
Db 493 CTTGTATGGCACTAACGCGACTACTACTCTCCAGGTAGAGAAATTTTAAATGGATG 552

Qy 1440 TTGCGCGCAGCAGAGAGTATTTCTATGATGTGGCTTTTGGGCAAGCAATAGCTCT 1499
Db 553 CTCAGAGCGGCTGGAAGAATATTTCTATGAATTTAGGTTTCTGGCTAAAGGTAAACGTTCT 612

Qy 1500 AGCAAAAAAACAATTTGTAGCAAGTAGAAGCGGCGCATTTGGTTTAAATTTGCATGAA 1559
Db 613 AACGATGCGAGCTTAGCCGATCAAAATTTGAAGCGGTGCGATTTGGCTTTAAATTTACGAA 672

Qy 1560 GACTCGGGCACAACCAAGTGCATCTGCTTGAAGCGTGGCAGATGAATAGCAT 1619
Db 673 GACTGGGCAACCACTCTTCTGCAATCAATCATGCTTAGATGTTTGGCGACAATAGCAT 732

Qy 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCTTA 1679
Db 733 GTGCAAGTGCCTATCCACACAGACACTTTTGAATGAAGCGGTTGTGTAGAAGACACTATG 792

Qy 1680 AATGCAATGAACGCGCGCCCATCATGCTACACATTTGAGGAGCGGTGGAGACAC 1739
Db 793 GCTGCTATTGCTGGAGCGCACTATGCACACTTTCCACACTGAAGCGCTGGCGGCGGACAC 852

Qy 1740 TCACCTGATTTATCACCATGGCAGCGAGCTCAATATTTCTACCTTCTCCACACCCCCC 1799
Db 853 GCTCCTGATATTTAAAGTAGCCGCTGAACACACATTTCTTCCCGCTTCCACTAACCCC 912

Qy 1800 ACTATTCCTATACCATTAATACGGTTGAGAACACTTATAGACATGCTCATGACATGCCAC 1859
Db 913 ACCATCCCTTTTCAACGCTGAATACAGAAAGCAGAGACATGAGACATGCTTATGTGTGTCAC 972

Qy 1860 CACCTAGACAAACGCGATCCGCGAGGATTTACAAATTTTCTCAAAAGCGGTATCCGCCCGGC 1919
Db 973 CACTTGGATTAAGCAATTAAGAAAGATGTTCAAGTTTTCAGTTCAGAGATCCGCCCTCAA 1032

Qy 1920 TCTATCGCGCTGAAGATGCTTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGCAT 1979
Db 1033 ACCATTCGCGCTGAAGACACTTTGATGACATGGGATTTTCTCAATCACAGTCTGTGAC 1092

Qy 1980 TCGCAAGCAATGGGCGTGACAGCGAAGTATTCCTGAACTTTGGCAGACTGCGGATTAAG 2039
Db 1093 TCTCAAGCGATGGGCGTGCGGTGAAGTTATCACTAGAACTTTGGCAAAACAGCTGACAAA 1152

Qy 2040 AATAAAGATTTGGTAAGCTTCTGAGATGCGCAAGATGAACGATATTTTCCGATTT 2099
Db 1153 AACAAGAAAGATTTGGCGCTTGAAGAAAGAAAGGCGATAACGCAACTTCAGATC 1212

Qy 2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTCACCCACGCGTGAGCGAGTAT 2159
Db 1213 AAACGCTACTGCTTAATATACACTTAACCCAGCGATCGCTCATGGGACTACGAGTAT 1272

Qy 2160 APTCGCTCTGTGGAAGAGGGCAAGATCGCGACTTGGTGGTGTGGAATCTGCTTTT 2219
Db 1273 GTCGGTCTGTAGAAAGTGGGCAAGTAGCTGACTTGGTATTTGGAGTCCAGCATTTCTT 1332

Qy 2220 GCGCTAAACCCAAATCGTATCAAGCGGTATGGTGTCTTCTCTGAAATGGGGCAT 2279
Db 1333 GCGCTGAAACCTTAACATGATCAATCAAGGTGGGTTCATTGCAATTAAGCCAAATGGGCGAT 1392

Qy 2280 TCTAACGCTGTGCGCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCAATCAC 2339
Db 1393 GCGAAGCTTCTATCCCTACCCCTCAACCGGTTTATTAAGAGAAATGTTGCTCATCT 1452

Qy 2340 GCGAAGGGGAAATTTGACACAGCATCACTTTTGTTCGAAAGTCCGCTATGAAATGGC 2399
Db 1453 GGTAAAGCTTAATACGATGCAAAACATCACTTTTGTCTCAAGCGCTTATGACAAAGGC 1512

Qy 2400 GTGAAGAAAGCTGGCTTAGCGGCTAAGTCTTACCGGTCAAAAATGCGCGCTTAACATC 2459
Db 1513 ATTAAGAAAGAAATAGGACTTGAAGACAAAGTGTTCGCGGTAAAAAATTTGCAAAAATC 1572

Qy 2460 ACCAAGAAAGCTTCAAGTTTCAACGCAAAACGCGAAATCAACCGTTCGATCCGAAACC 2519
Db 1573 ACCAAAAAGACATGCAATTTCAACGACACTACCGCTCAATTTGAAGTCAATCTCTGAAACT 1632

Qy 2520 TTCGAGGTCTTTGTAGATGGCAAACTCTGACACTCTTAAACCCACCTCGAAGTGCCTCTA 2579

Db 1633 TACCATGTTCTGGATGGCAAGAGTAACTTCTTAACCAAGCTAATAAAGTAGCTTG 1692
Qy 2580 GCCAGCGCTACACTTTCTTAG 2603
Db 1693 GCGCAACTCTTTAGCATTTTCTAG 1716

RESULT 14
US-09-402-100-1
; Sequence 1, Application US/09402100
; Patent No. US20010019834A1
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helio
; FILE REFERENCE: 0136/OG140
; CURRENT APPLICATION NUMBER: US/09/402,100
; EARLIER FILING DATE: 1999-09-27
; EARLIER FILING NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: ()..()
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-1

Query Match 29.9%; Score 860.6; DB 3; Length 2385;
Best Local Similarity 69.8%; Pred. No. 1.7e-178;
Matches 1169; Conservative 0; Mismatches 504; Indels 1; Gaps 1;

Qy 900 AAAATGAAAAAACAAGATATGTAATACCTACGAGCCACCAAGGCGATAAAGTGGC 959
Db 7 AAGATTAGCAGAAAGAAATATGCTTCTATGATGGCCCTACTACAGCGATTAAGTAGA 66
Qy 960 TTAGGAGATACCGATCTTTGGGAGAGAGTAAAGATGATACACCTATGGCGAAGAA 1019
Db 67 TTGGCGATACAGACTTTGATCGCTGAAGTAGAATGACTACACCAATTTATGGTGAAGAG 126
Qy 1020 CTTAAATTTGGCGGCTAAACCTATCCGTGAGGATGAGGTGAGAGCAATAGCCCTGAT 1079
Db 127 CTTAAATTTGGCGGCTAAACCTATGAGAGAGGATGAGCAATTAACACACCTAGC 186
Qy 1080 GAAAAACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
Db 187 AAAGAAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 246
Qy 1140 AAAGCCGACATTTGGATTTAAACCGCAAAATCCATGGCATTTGGCAAGGAGGAGAAACAG 1199
Db 247 AAAGCGATATTTGATTTAAAGATGAGCAAAATCGCTGGCATTTGGTAAAGGCGGTAAACAA 306
Qy 1200 GACATGCAAGATGCGGTAAAGCCCTCATATGATGATGATGATGATGATGATGATGATGAT 1259
Db 307 GACACGCAAGATGCGGTAAACCAATCTTAGCGTGGGTCTTGCTACTGAAGCCTTAGCC 366
Qy 1260 GGGGAGGATGATTTATACGCTGGGGGATGATGATGATGATGATGATGATGATGATGATGAT 1319
Db 367 GGTGAAGGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
Qy 1320 CAACAATTTCCCTACCGCTCTAGCAATGGCGTTTACACCAATGTTTGGAGGCGGACAGGT 1379

Db 427 CAACAAATCCCTACAGCTTTTGAACGGTGTAAACCAATGATTTGGTGGCGAAGTGGC 486
Qy 1380 CCGTGTAGATGGCAAGATGCGACTACTATCCTCCGGGCAAAATGGAATTTGCAACCGCATG 1439
Db 487 CCGTGTAGTGGCACTAAACCAACCACTATCCTCCAGGTAGAGAAATTTAAATTTCAATG 546
Qy 1440 TTGGCGGAGAGAGATTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
Db 547 CTCAGAGCGGCTGAAGATTTCTATGAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 606
Qy 1500 AGCAAAAAACAATTTGTAGAACAAAGTAGAGCGGCGCATTTGGTGGTGGTGGTGGTGGTGGT 1559
Db 607 AACGATGCAAGCTTAGCCGATCAAAATTTGAAGCTGGTGGCAATTTGGCTTTAAATCCAGCA 666
Qy 1560 GACTGGGCAACCAACCAAGTGGATCGATCACTGCTTGAGCGGTGGCAGATGATGATGATGATGAT 1619
Db 667 GACTGGGCAACCAACCACTCTCTGCAATCAATCATGCGTTAGATGTTGGGAGCAATAGCAT 726
Qy 1620 GTGCAAGTTTGTATCCACACCGCATACAGTCAATGAGGAGGTTTATGATGATGATGATGATGAT 1679
Db 727 GTGCAAGTCTGCTATCCACACAGACACTTTGAATGAAGCGGTTGCGTGGAGACACTATG 786
Qy 1680 AATGCAATGAGGCGGCGCATCCATGCTTACCACTTGAAGCGGTTGCGTGGAGGAGGAGGAGC 1739
Db 787 GCGACTATTGGCGGAGCGCATATGACACTTACCACTGAAGCGGCTGGCGGCGGAGC 846
Qy 1740 TCACCTGATTTATCAACATGGCGAGCTCAATATTTCTACCTCTCTCCACACACCCCT 1799
Db 847 GCTCTGATTTATTAAGTGGCGGAGTGAACACACATCTTACCTGCTTCCACTTAACCC 906
Qy 1800 ACTATTCCTATACCAATTAATACGTTGAGAACACTTAGACATGCTATGATGATGATGATGATGAT 1859
Db 907 ACTATTCCTTTTCCCGTGAATACAGAGCGGAGCACTGATGATGATGATGATGATGATGATGATGAT 966
Qy 1860 CACTAGACAAACGCTACCGGAGATTTCAATTTTCTCAAGCGGATTCGCGCCCGG 1919
Db 967 CACTTGGATAAAGCAATTAAGAGATGTCCAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1026
Qy 1920 TCTATCGCGCTGAAGATGCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1979
Db 1027 ACCATTTGGCGTGAAGACACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
Qy 1980 TCGAAGCAATGGGCGTGGCGAGGAGTGTCTCGAATTTGGCAGACTGGCGGATAAG 2039
Db 1087 TCTCAAGCGATGGCGGTGGTGGTGAATTACTAGAACTTTGCAACAGCTGACAA 1146
Qy 2040 AATAAAGAAATTTGGTGAAGCTTCTGAAGATGGCAAGATAACGATTAATTTCCGCAAT 2099
Db 1147 AATAAAGAAATTTGGCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206
Qy 2100 AAGCGTACATCTCAAAATACACTATCAACCGCTTTGACCCAGCGCTGAGCGAGTAT 2159
Db 1207 AAGCGTACATCTTAAATACACCAATTAACCGAGATGCTCATGGGATGAGCGAGTAT 1266
Qy 2160 ATCGGCTCTGGAGAGGCGAAGATCGCGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2219
Db 1267 GTCGGTTCTGAGAGTGGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1326
Qy 2220 GCGGTAAACCAAAATCGTATCAAGCGCGTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2279
Db 1327 GGTGTGAACCAACCATGATCATCAAGCGCGGTTTATCGCATTTGATGATGATGATGATGATGATGAT 1386
Qy 2280 TCTAAGCGGCTGTGGCCACTCCCAACCGGTTTATTAACCGGAAATGTTTGGGATCAC 2339
Db 1387 GCGAAGCGCTTCTATCCCTTACCCCAACCAAGTGTATTAACAGAGAAATGTTTGGCTCATCAT 1446
Qy 2340 GCGAAGCGAAATTTGACACCAAGCATCTTTTGTTCCTCAAAAGTGGCTTATGAAATGGC 2399
Db 1447 GGTAAAGCTAAATAGATGCAAAACATCACTTTTGTCTCAAGCGGCTTATGCAAGAGC 1506
Qy 2400 GTGAAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAAGTGGCTTAAATC 2459

Db	563922	TATCAAGCCGATATTGGTATTCTGATGGCGGTATTGTGGGTATTGGACAGAGTAA	563863
Qy	1196	CAAGGACATGCAAGATGCGTAAGCCCTCATATAGTCTGCTGGGTGGCGACAGAGCACT	1255
Db	563862	CCCTGACACCATGATGATACGTCACACCAATATGATTTATCGGTCAAGCAGCGAATTCA	563803
Qy	1256	AGCAGGGGAAGGTATGATTATACCGCTGGGGGAATCGATTTCACACACCCACTTCTTTC	1315
Db	563802	TAAACGGTGCAATTTAATTTGCAACCGCTGGGTATCGATACCCCATTCATTTATTGG	563743
Qy	1316	TCCACAACAAATTCCTACCGCTTAGCCCAATGGCGTTTACCAACATGTTTGGAGGGCGCAC	1375
Db	563742	TCCACAACAGCAACAAATGCAATGCAATTTGAAGTGGCTTACCAAGTTAAATTTGGTGGGAAC	563683
Qy	1376	AGTCTCTGTAGATGGCAGCAATGGCACTATCATCTCGGGCAAAATGGAATTCGACCG	1435
Db	563682	TGGCCCTGCTGATGGTACACACGCAACCACTTGTACCCCTGGCGCATGGTATATGGAACG	563623
Qy	1436	CATGTTGGCCGACGACAGAGATTTCTATGAATGTGGCTTTTGGGCAAAAGGCAATAG	1495
Db	563622	TATGTTTCAACGGCGAGAAGCTTTGCCGGTAAACGTCGGATTTTTTGGTAAAGGCAACTG	563563
Qy	1496	CTCTAGCAAAAAACAACTTTGTAGAACAAAGTAGAGCGGGCGGATTTGGTTTAAATTTGCA	1555
Db	563562	TTCAACCTAGATCTCTGCTGGTGAGCAAAATGGAACGGGTGCATTTAGGTTTAAAAATCCA	563503
Qy	1556	TGAAGACTGGGGCAACACCAAGTGGCGATCGATCACTGTTGAGCGTGGCGAGATGAATA	1615
Db	563502	CGAAGACTGGGGTGCAACCGCTGCCGTGATTGATTCGCTTTAAAGTAGCAGATGAAT	563443
Qy	1616	CGATGTGCAAGTTTGTATTCACACCGATACAGTCAATGAGGAGGTTTATGTAGATGACAC	1675
Db	563442	GGATATTTCAAGTGGCCATTTCACACAGACACGCTAAATGAAAGTGGCTTTTGGAAAGACAC	563383
Qy	1676	CCTAAATGCAATGACGGCGCGCATCCATCCCTACCACATTTGAGGGAGCGGGTGAGG	1735
Db	563382	GATGAAAGCAATGATGACGAGTCAITTCATCTTTCCATGAGGGCGAGGTGGTG	563323
Qy	1736	ACACTCACTGATGTTATCAGCATGGCAGGGAGCTCAATATTTCTACCCCTCTCCACCAC	1795
Db	563322	TCATGCACTGACATCAATTAAGCAGCGATGATTTCAACGATTTACCTGCTTCAACCA	563263
Qy	1796	CCCACTATTCCTTATACATTAATACGGTTGCGAGAACACTTAGACATGCTCATGACATG	1855
Db	563262	CCCGACTCGTCCGTTTACCAAAAAACCACTGATGAACATTTGGATATGTTGATGGTTG	563203
Qy	1856	CCACCACTAGACAAACGATCCGCGAGATTACAAATTTTCTCAAAGCGGTATCGGCC	1915
Db	563202	CCATCACTTAGATAAACCGGTGCGCGAAGACGTAGCTTTTGGCGGATAGCCGATCCGCC	563143
Qy	1916	CGGCTCTATCGCGCTGAAGATGTCTCCATGATATGGTGTGATCGGATGACAAAGCTC	1975
Db	563142	TGAACCAATTTGACGACAGAAGATATTTTGCATGATATGGCGTCTTCTCAATTATGAGTTC	563083
Qy	1976	GGATTGCGAAGCAATGGGGCGTGCAGGCGAAGTATCTCGAACTTGGCAGACTGCGGA	2035
Db	563082	AGACTCTCAAGCGATGGGACGTATTGGCGAAGTCGTTATTCGTATCATGGCAAACTGCAGA	563023
Qy	2036	TAAAGATAAAAAAGNATTTGGTAAGCTTCCGAAGATGCGAAGATAAGCATATTTCCG	2095
Db	563022	TAAAGTAAAAATGCAACGTGGTGAGCT-----AGGTAATGAAGGAAACGATAAATTCG	562969
Qy	2096	CATTAAGCGCTACATCTCCAAATACATATCAACCCGCTTTGACCCAGCGGTGAGCGA	2155
Db	562968	TATTAACGATATATTCGGAATAACCATCAACCCAGCAATTGACATGGTATTTGCGGA	562909
Qy	2156	GTATATCGGCTCTGTGGAAGGGGCAAGATCGCGACTTTGGTGGTGTGGAATCTTGCCTT	2215
Db	562908	GCATATTGCTCGTTAGAAAGTGGTAAATCGCAGATATCGTGTATGGAACCGATGTT	562849
Qy	2216	TTTTGGCGTAAACCACAAATCGTATCAAGCGGATATGGTGGTCTTCTCTGAAATGGG	2275
Db	562848	CTTTGGCGTAAACCTGAAGTGGTGTATTAATAAAGGCTTTATTAGCTATGCGAAAAATGGG	562789

Search completed: August 10, 2006, 03:37:28
Job time : 4674 secs

Qy	2276	CGATTCTAACGGCTGTGTGCCCACTCCCCCAACGGGTTTATTACCGGAAATGTTTGGGCA	2335
Db	562788	CGATCCAAATGCCTCAATTTCCAACACCGCAACCTGTATTCTACCGTCCAATGTACGGTGC	562729
Qy	2336	TCACGGCAAGGCGAAATTTGACACACAGCATCACTTTTGTTCACAAAGTCGCCCTATGAAA	2395
Db	562728	ACAAGGCTTAGCAACCGCAACAAACAGCAGTATCTTTGTTCACAAAGCGCTGAAAAGC	562669
Qy	2396	TGGCGTGAAGAAAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAACTGCCGTAA	2455
Db	562668	TGATATTCTGTCGAAGTTTCGGTTTACACAAAGAACCAATTGCTGTGAAGGCTGCCGCA	562609
Qy	2456	CATCACCAAGAAAGACTTCAAGTTCAACGACAAACGCGCAAAATCACCGTCGATCCGAA	2515
Db	562608	CGTAGGTAAAAAAGATCTGGTTTCATAATGATGTAAACACCAACATTTACTGTTGATGCTGA	562549
Qy	2516	AACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGC	2575
Db	562548	ACGTTATGAAGTTTCGAGTGGACGAGAGTTAATTACCTGTGAACCAAGTGGATAGCGTACC	562489
Qy	2576	TCTAGCCCGCGCTACACCTTTCTTCTA	2602
Db	562488	ATTGGGTACGCGATATTTCTTATCTA	562462

[illegible]

Db	27933	AAGTTGCAATTAGAAAGTGGTATCACAAACATATCGGTGGCGTGACAGGCGGCTCTGAGG	27999
Qy	1390	GCACGAATCGCGACTACTATCATCTCCGGGCAAAATGGAACTTGGCACCGCATGTTGGCGCGAG	1449
Db	27993	GTACGAGGCGCAACTGTACACCTGGACCATGGCACTTTACATCGTATGTACTCGAG	28052
Qy	1450	CAGAAAGATTTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAC	1509
Db	28053	CAGAGTCATTCGGTTGAAATATGGCTTTACTGCTAAAGGCAAGCTGTAACACACACAG	28112
Qy	1510	ACTTTGTAGAACAAAGTAGAAGCGGGCGGATGGTTTAAATTTGCATGAAGACTGGGGCA	1569
Db	28113	CTTTAGTGGAAACAAATTTCACTGGTGGCAATCGGCTTAAAGTGCATGAAGACTGGGGTG	28172
Qy	1570	CAACCAAGTGCAGTCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTT	1629
Db	28173	CAACGCTTCAGCATTTGATCATGCTTTTACAGTCCGAGATGACTATGACGTACAAATCG	28232
Qy	1630	GTATCCACACCGATACAGTCAATAGGCGAGGTTATGTAGATGACACCTAAATGCAATGA	1689
Db	28233	CGTTGCACGCAGATACGTTAAATGAAGCTGGTTTCTTGGAAAGAAACAATGGCTGCAGTAA	28292
Qy	1690	ACGGGCGCCATCCATGCTTACCACATTTGAGGAGCGGGTGGAGCACTCACTCATGTG	1749
Db	28293	AAGATCGCGTATTGCAATATGTATCATACGGAAGGTGTGGTGGTGGTTCATGCGGCCGATT	28352
Qy	1750	TTATCACCATGCGAGCGAGCTCAATATTTCTACCCCTCTCCACACCCCCCACTATTCCT	1809
Db	28353	TAATTTAAGTCTGCTGGTATGCTAACTTTTACCTTCTTCAAACAATCCGACGTACCTT	28412
Qy	1810	ATACCATTAAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCACACCTAGACA	1869
Db	28413	ATACTGTAATACAAATCCGATGAACATTTAGATATGGTTATGATTAACCATCATTTAAATG	28472
Qy	1870	AACGCATCGCGAGGATTTACAATTTTCTCAAGCCGATATCCGCCCGGCTCTATCGGG	1929
Db	28473	CATCTATCCCTGAAGATATTCGGTTTTCGGATTCAGCTATTCGTAAGAAACCAATTCGAG	28532
Qy	1930	CTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAA	1989
Db	28533	CTGAAGATGTACTTCAAGATATGGGCGTATTTAGTATGGTAACTTCTGATTTCAACGCA	28592
Qy	1990	TGGGCGTGCAGGCGAAGTGATTCCTCGAACTTTGGCAGACTCGGATPAAGATAAAAAG	2049
Db	28593	TGGGACGTGTCGAGAAGTAAATTAACCGTACTTTGGCAAGTTGCTCACCGTATGAAGAAC	28652
Qy	2050	AAATTTGGTAACTTCTGAAAGTGGCAAGATAAAGATAATTTCCCGCATTTAAGCGCTACA	2109
Db	28653	AACGTGGATTTATAGATGGCGACAGTGAATACAATGACAAATAATTCGTTTAAACGCTATA	28712
Qy	2110	TCCTCAAAATACACTATCAACCCCGCTTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTG	2169
Db	28713	TAGCAAAATATACAATTAACCCAGCCATTTACATGATATTTCTGACTATGTAGGTTCAA	28772
Qy	2170	TGGAAGAGGCGAAGATCGCCGACTTGGTGGTGTGGAATCTGCTCTTTTGGCGTAAAAAC	2229
Db	28773	TTGATGAAGGTAAATTTAGCCGACATCATCTTTTGGGAACCAGCATTTCTTGGCGTTAAAC	28832
Qy	2230	CCAAATTCGTGATCAAAAGCGGTATGGTGTCTTCTCTGAAATGGGCGATTTCTAACCGGT	2289
Db	28833	CTGATGTCAATCGTTAAAGCGGGTTAATCAACGCTGCAATCAACGGAGATGCGAATGGCT	28892
Qy	2290	CTGTGCCCACTCCCAACCGGTTTATTTACCGCGAAATGTTTGGGCATCACGGCAAGCGCA	2349
Db	28893	CTATCCCTACTTCAGAACCTTTTAAATATCGCAAAATGTATGTTCAATTAGTTCGAAATC	28952
Qy	2350	AAATTTGACACCAAGCATCACTTTTGTGTTTCCAAAGTCGCTATGAAATTTGGCGTGAAGAAA	2409
Db	28953	TACAAAGTACATCGATGACTTTTGTGTTCTACAACTGCTTATGAAACGATATTCGTAAC	29012
Qy	2410	AGTGGGCTTAGAGCGCCAGTTCTTACCGTCAAAAACCTCCCGTAAACATCAACCAAGAAAG	2469
Db	29013	TTTTTAGGCTTAAACCGCAATTAAGACCTGTGCACATATCCGTAAATTAAGTAAAAAAG	29072

Qy	2470	ACTTCAAGTTC	AACGACG	CAAAACGG	CCAAAAAT	CACCGT	CGATCCG	AAAAACCTT	CGAGGTCT	2529
Db	29073	ATATGAAAAA	CAATATG	CAACACCA	CAGATTT	TAGACGTT	GACCA	CAAAAT	GAAAGTTT	29132
Qy	2530	TTGTAGATG	CGAAACT	CTGCACCT	CTTAAC	CCACCT	CGCAAGT	GTGCTCT	AGCCAGCGCT	2589
Db	29133	TTGTTGATG	GAGAGAAA	TTTAC	AGCGAA	CCGTCT	ACAGATT	ACCATTT	AACACACGCT	29192
Qy	2590	ACATTTCT	TCTTAGG	2604						
Db	29193	ATTTCCT	TATTTAGG	29207						

RESULT 2

```

US-11-348-413-11112
; Sequence 11112, Application US/11348413
; Publication NO. US20060160121A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES

```

```

// DIFFERENTIAL AMPLIFICATION PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
// TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
// FILE REFERENCE: 031896-084100 (AM 101724)
// CURRENT APPLICATION NUMBER: US/11/348,413
// PRIOR FILING DATE: 2006-02-07
// PRIOR APPLICATION NUMBER: PCT/US05/035471
// PRIOR FILING DATE: 2005-10-05
// PRIOR APPLICATION NUMBER: US 11/243,445
// PRIOR FILING DATE: 2005-10-05
// PRIOR APPLICATION NUMBER: US 60/615,573
// PRIOR FILING DATE: 2004-10-05
// NUMBER OF SEQ ID NOS: 1276209
// SEQ ID NO 1112
// LENGTH: 1716
// TYPE: DNA
// ORGANISM: Staphylococcus epidermidis
// FEATURE:
// NAME/KEY: misc feature
// LOCATION: (1)..(1716)
// OTHER INFORMATION: WAN01UFW7; ORF: 00000000011100; Cluster contains WAN01OSH0
// OTHER INFORMATION: :SEI863:ureC:urease alpha subunit:ATCC12228:NC_004461.1
// US..11-348-413-11112

```

	Query Match	19.7%; Score 566.8; DB 9;	Length 1716;
	Best Local Similarity	59.4%;	Pred. No. 6e-124;
	Matches 1018;	Conservative 10;	Mismatches 672; Indels 15; Gaps 3;
QY	898	TGAAATGAAAAACAAGATATGTAAATACCTTACGGACCACCAAGGCCGATAAAGTCG	957
DB	8	TTAAAATGACACAATCTCAATACACAAGCTTTATTGGACCAAATGTGAGGAGACTCTGTGA	67
QY	958	GCTTTAGGAGATCCGATCTTTTGGCGAGAAGTAGAACATGACTATACCACCTATGCSCGAAG	1017
DB	68	GATTAGGAGNATCGAACTTGTTTTGCACAGTTGARAAGACTATGCNAATTATGGAGATG	127
QY	1018	AAC TTAAATTTGCGCGGGTAAAAACTATCCGTGAGGGTATGGGTCAGA----GCAAT-	1070
DB	128	AAGCTACTTTTCGGTGGCGGAAAAATCAA T TC GTGATGGTATGGCTCAAAATCCTTAATGTGA	187
QY	1071	--AGCCCTCATGAAACACCCCTAGATTTAGTCATCACTAACGCGATGNTTATCGACTACA	1128
DB	188	CMAAGAGATGATAAAAATGTAGCCGNTTTAGTTTAACTAACGCATTAA'TPATTGATTTATG	247
QY	1129	COGGGATTTTCAAAGCCGACATTTGGGATTTAAAAACGCAAAAATCCATGGCATTGGCCAAG	1188
DB	248	ACAAGATTGTTAAGACAGAYATCGGAATTA AAAAATGGTTATATTTTTTAAGATCGGTAAG	307
QY	1189	CAGGAAACAAAGCACATGCAAGATGGCGTAAAGCCCTCATATPGTGTGGGTGGGCACAG	1248
DB	308	CT'GGAAACCCAGATAAATGGGTAACGTTTT-----TGACATCATCTTTGGTGCACAACATG	361

Qy	1249	AAGC	ATAGCAGGGG	AAGTATGATATATACCGCTGGGGGAATCGATTTCACACACCCACT	1308
Db	362	ATATTA	TGTGCTGAAAGG	TAATAATTTGTTACTGCGGCGGTATCGATAACACACGTGCACT	421
Qy	1309	TCCTTT	CTCCACAA	CAATCCCTACCGCTTAGCCAATGGCGTTACAAACATGTTTGGAG	1368
Db	422	TYATCA	ATCCTGAA	CAAGCTGAAGTTGCATTTGAGATGGTATTTACAAACGATATTCGGTG	481
Qy	1369	CGGC	CACAGGTCCTGTAGAT	GGCACGAATGCGACTACTCATCTCCGGGCAAAATGGAAC	1428
Db	482	GAGGA	CTGGTGCTTCTGAA	GGTGCTTAAGCGACTACTGTAAACACAGGACCTTTGGCATA	541
Qy	1429	TGCA	CCGATGTTGGCG	CGAGCAGAGATATTTCTATGAATGTGGGCTTTTTTGGGCAAG	1488
Db	542	TTCA	TCGATGTTAGA	AGCAGCAGAGAGATGCCATATTAATGTAGGATTTACTCGTAAAG	601
Qy	1489	GCA	TAGCTTAGCA	AAAAACACTTGTAGAACAAAGTAGAAGCGGGCGGATTCGTTTA	1548
Db	602	GKAA	GCTGTCAATCAT	ACTGCACTTAATGAACAAATTCATCGAGGCGCTATAGGCTTTA	661
Qy	1549	AA	TGCA	TGAAGACTGGGCGACAAACCAAGTGCATCGATCTGTCTGAGCGGTGGCAG	1608
Db	662	ANG	TACATGA	AGATTGGGAGCTACACCTTCAGCATTAAGTCATGACGTTCGAG	721
Qy	1609	ATGA	ATACGATGTG	CAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTATGTAG	1668
Db	722	ATGA	TTTGATRTTCA	AGTCGCTTTACATGCAGAYACATTAATGAAGCTGGATTTATGG	781
Qy	1669	ATGA	CCCCCTAA	ATGCAATGAACGGGCGCGCATCGATCGCTTACCACATTTGAGGGAGGG	1728
Db	782	AAGA	TACAAT	TGGCTGTTAAGAAGATCGTGATTTGCCATATGATATCATCTGAAGGAGCTG	841
Qy	1729	GTG	GAGGACACTCA	CCCTGATTTATCACCATGGCAGGCGAGCTCAATATTTCTACCCCTCCT	1788
Db	842	GTGG	TGTGTAC	GCCTGACTTAATCAATCAGCTGCATATTCAAACATCTTACCTTCTT	901
Qy	1789	CCAC	CCCCCCTATTTCC	CTATACCAATTAATACGGTTGACAGAACACTTAGACATGCTCA	1848
Db	902	CTACA	ACC	CAACATTAACCTTTACACACAACTGTAGATGAACATTTAGACATGGTTA	961
Qy	1849	TGAC	ATGCCACG	ACTAGACAAAGGCATCGCGAGGATTTTCAATTTTCTCAAAGCCGTA	1908
Db	962	TGAT	YACTCC	CACTTTAATGCTTCAATACAGAAGACATTTGCATTTGCGAGATTTCTCGTA	1021
Qy	1909	TCG	CCCGGCTCTATCG	CGGCTGAAGATGTGCTCCATGATATGGGTGTCATCGCGATGA	1968
Db	1022	TAC	TTAAGGA	AACTATAGCAGCAGAGAYGATTTACAAATATGGGCGTATTTAGTATGG	1081
Qy	1969	CAAG	CTCGGATTCG	CAAGCAATGGGCGTGACGCGAAGTATTTCTCGAATTTGGCAGA	2028
Db	1082	TAAG	TTCAGATTCACA	AGCAATGGGACGTGCTGCTGAAGTTGTAAACGCTACTTTGGCAAG	1141
Qy	2029	CTG	CGGATTAAG	ATATAAAGAAATTTTGTAAGCTTCTGAAAGTAGGCAAGATAACGATA	2088
Db	1142	TTG	CACACCGTATGA	AAAGAACCAACGGGCGACATAGATGCTGATTTTGAATATCACGATA	1201
Qy	2089	ATTT	TCGCAATTAAG	CGCTACATCTCCAAATACACTATCAACCCCGCTTTTCACCCACGGCG	2148
Db	1202	ATA	TCTGATTAAC	CGTTTACATTCGCAAAATATACATCAATCTCGCAATTTACACATGGTA	1261
Qy	2149	TGAG	CGAGTATATCGG	CTCTGTGGAAGAGCGCAAGATCGCCGACTTTGGTGGTGGGAATC	2208
Db	1262	TTTCT	GACTATGTTGG	ATCTGTGAGCGGGTAACTTTGCCGATTTAGTAAATGTGGGAAC	1321
Qy	2209	CTG	CGTTTTTTTGG	CGTAAMACCAAAATCGTGATTCAAAGCGGATATGGTGGTCTTCTCTG	2268
Db	1322	CAGA	ATTTCTCGGTG	CCAAACCCGATCTTGTGTTTAAAGGTGGCATGATTTAACTC	1381
Qy	2269	AAAT	TGGGATTTCA	ACGGTCTGTGCCACTCCCAACCGGTTTATTACCGGAAATGT	2328
Db	1382	TAA	TGTGTATGT	TAATGGCTCCATACCAACATCAGAGCCTTTTGAATATTCGCAAAATGT	1441
Qy	2329	TTGG	GCATCAGG	CAAGGCGAAATTTTGACACCAAGCATCACTTTTGTTC	2388

```
Db 1442 ATGGTCAATTTGGTGGTAACATTTACATATCTGCTATGACTTTTGTCTTAAACACTGCAT 1501
Qy 2389 ATCAAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAAC 2448
Db 1502 ATGAATAACGGTATTTATCGTCACTCAATCTAAACGATGGTTGACCAAGTTAGAATA 1561
Qy 2449 GCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAACCGCAAAAATCACCGTCG 2508
Db 1562 TTAGAAATTTAACTAAGGCAGATATGAAAAATAAATATGCTACACCTTAAATAAGATGAG 1621
Qy 2509 ATCCGAAACCTTCGAGGCTTTGTAGATCGCAACTCTGCACCTCTAAACCCACCTCGC 2568
Db 1622 ATCCACAAACATATGAGGTATTCGTTGATGGTAAATAAAATCACAAGTGAAGCAGACAG 1681
Qy 2569 AAGTGCTCTAGCCAGCGCTACACTTTCTTCTAG 2603
Db 1682 AATTACCWTTAACAAAGATACTTCTTATCTAG 1716

RESULT 3
US-10-471-571A-3299
; Sequence 3299, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3299
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3299

Query Match 18.7%; Score 539.6; DB 6; Length 1713;
Best Local Similarity 58.9%; Pred. No. 1.16e-117;
Matches 1010; Conservative 0; Mismatches 689; Indels 15; Gaps 4;

Qy 896 CATGAAATCAAAAACAAGATATGTAAATACCTACGACGCCACCAACGCGATAAAGT 955
Db 6 CTTTAAATGACGCAAAATCAATATACAGCTTATACGGTCCAACTGTTGAGATTCAT 65

Qy 956 GCGCTTAGGAGATACCGATCTTTGGGAGAGTAGAATCATGATATACCACTATGGCGA 1015
Db 66 TCGTTTAGGTGATACGATCTATTGCTCAATAGAAAAGACTATGGGTTTATGGTGA 125

Qy 1016 AGAACTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGTCA-----GAGC 1067
Db 126 AGAAGCTACTTTTGGTGGTAAATCTATTAGAGACGGTATGGCGAAAAATCCTCGTGT 185

Qy 1068 AATA-GCCCTGATGAACAACCCCTAGATTAGTTAGTTCATCACTAAACCGGATGATTACGACTA 1126
Db 186 AACACGTGATGACGTGAACGTTGACAGACCTTGTCATTTCTTAATGCCGTTATTATCGATTA 245

Qy 1127 CACCGGATTTACAAAGCCGACATTTGGATTGGAATTAATAACGCAAAATCCATGGCATTTGCCAA 1186
Db 246 CGATAAGTGGTTAAAGCTGATATAGGCAATTAATAATGGTTATATTTCGCCATAGGTAA 305

Qy 1187 GGCAGGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGGTCGTGGGTGGGCAAC 1246
Db 306 TGCAGGCAACCCAGATAT----AATGATAATGTC--GACATTATTAGGTTCAACAAC 359

Qy 1247 AGAAGCACTAGCGGGAAGGATGATATTATTACCGCTGGGGGAATCGATTACACACCCA 1306
Db 360 AGATATCAITGGCCCTGAAGGTAATAATCGTCACTGCTGGTGGTATTGATACTCATTTCA 419

Qy 1307 CTTCCCTTCTCCACACAATTTCCCTCCTACCGCTCTAGCCAAATGGCGGTTTACCAACATGTTGG 1366
```

```
Db 420 TTTTATTAAATCTGAAACAAGCAGAGGTGCGATTTAGAAAGTGGTATTACGACTCATATTGG 479
Qy 1367 AGCGGCAAGGTCTCTGTAGATGGCACGAATGCGACTACTACTCATCTCCGGGCAAAATGGAA 1426
Db 480 TGGTGGTACTGGTCTTCAGAAAGTTCTTAAGCAACAACCTGTAACTCCAGGTCCTATGGCA 539
Qy 1427 CTTGCACCGCATGTCGCGCAGCAGAAAGATTTCTATGAATGTGGGCTTTTGGGGCAA 1486
Db 540 TATTATAGAAATTTAGAAAGCTGCCGAAGTTTACCGATTAATGTTCGGTTTACAGGTAA 599
Qy 1487 AGGCAATAGCTCTAGCAAAAACAACCTTTGTAGAACAAAGTAGAAGCGGCGCATTTGGTTT 1546
Db 600 AGGCAAGCAACAATCCAACTGCACTCATTTGAACAAAATCAATGCGGAGCAATTGGATT 659
Qy 1547 TAAATTCATGAAGACTGGGCGACAAACCAAGTGCATTCGATCACTGCTTCGAGCGTGGC 1606
Db 660 AAAAGTATGAGAGACTGGGGTGCACACCATCTCTCTTTGAGTCTGATCATGATTAGATTGG 719
Qy 1607 AGATGAATACGATGTGCAAGTTTGTATCCACAACGATACAGTCAATAGAGCGAGTTTATGT 1666
Db 720 TGATGAATTTGATGTTCAAAATTGCAATTACATGCAGATACTTTTAAATGAAGCAGGATTTAT 779
Qy 1667 AGATGACACCCCTAAATGCAATGAACGGGCGCGCATCCATGCTACCACTTACGAGGAGC 1726
Db 780 GGAAGACACAATGGCTGCTGTAAAGACCGTGTACTTCAATATGTACCATACTGCAAGGTGC 839
Qy 1727 GGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGCGAGCTCAATATTTCTCAACCTC 1786
Db 840 TGGTGGCGGTCATGGGCTGATTTAAATCCGCTGCATTTTCAATATTTTACCTTTC 899
Qy 1787 CTCACACACCCCACTATTCCCTATACCATTAATACGGTTGCAGAACACTTACAGATGCT 1846
Db 900 ATCTAGAAATCCAACTTTGGCTTATACACATAATACTGTAGATGAACATTTAGATATGGT 959
Qy 1847 CATGACATGCCACCACTTAGCAAAACGATTCGCGAGGATTTACAAATTTTCTCAAGCCG 1906
Db 960 AATGATTACTCACCATTTTAAATGCGGCTATTCTGAAGATATCGCATTCGCAATTCACG 1019
Qy 1907 TATCGCCCGGCTCTATCGCGGCTGAAGATGCTGCTCCATGATATGGGTGTGATCGCAT 1966
Db 1020 TATTCGTAAAGAAAGATTTGCAGCAGAGATGTTCTGCAAGATATGGGTGATTCAGTAT 1079
Qy 1967 GACAAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGCAAGTGAATTCCTCGAACTTGGCA 2026
Db 1080 GATTAGTTCCGATTCACAAGCAATGGCGGTGTAGTGAAGTAAATTACACGACATGGCA 1139
Qy 2027 GACTCGGATTAAGATAAAAAAGAAATTTGGTAAGTTCCTGAAGATGGCAAGATAAACGA 2086
Db 1140 AGTAGACATCGCATGAAAGAAACAACGTTGTCCTTTAGATGGTGATTTTGAACATATGA 1199
Qy 2087 TAAATTCGCGATTAAGCGCTACATCTCCAAATACACTATCAACCCGCTTTGACCCACGG 2146
Db 1200 TAAATTCGCGATCAAAACGTTATATCGTAAATATACAAATTAACCCAGCAATTACACATGG 1259
Qy 2147 CGTAGCGAGTATATCGGCTCTGTGGAGAGGCGAAGATCGCGGACTTGGTGGTGGGAA 2206
Db 1260 TATTTCTGAATATGTAGGATCTATCGAGCGGGCAAACTAGCTGACATTTCTTATGGGA 1319
Qy 2207 TCCTGSCCTTTTGGCGTAAAAACCCAAATTCGTGATCAAAAGCGGATATGGTGGTCTTCTC 2266
Db 1320 CCCAATTTCTTTGGGGTTAAACCTGAAATTTAGTTGTAAAGGGCGGATTAATTAACCTCTGC 1379
Qy 2267 TGAATGGCGGATTTCAACGGCTCTGTGCCACTCCCAACCGGTTTATTTACCGCGAAAT 2326
Db 1380 CGTAAATGCGATGCAAAATGGTTCTTATACCTACATCTGAACCGGATGAAGTACCGTAAAA 1439
Qy 2327 GTTTGGGATCAACGCGCAAGGGGAAAATTTGACACACGACATCACTTTTGTTCCTCAAGTCGC 2386
Db 1440 GTATGGTCAATACGCGGGAACCTTACAAGTAGGTCAATGACATTTCTGTCTTAAACTGC 1499
Qy 2387 CTATGAAATGGCGTGAAGAAAAAGCTGGGCTTTAGAGCGCCAAAGTTCTTACCGGTCAAAA 2446
```

Db 1500 TTATGAAATGGTATCAATCGTGCATTAAATTTTAAACGCGATGGTGGCTCCAGTTTAAAA 1559
Qy 2447 CTGCGGTAAATCAACCAAGAAAGACTTCAAGTTTCAACGACAAACCGGCAAAATCAACCGT 2506
Db 1560 TATTAGACAATTTATCTAAAGCAGATATGAAAAATAACAGTGCACACCTAAATTAGAGCT 1619
Qy 2507 CGATCCGAACCTTCGAGGTCCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTC 2566
Db 1620 TGATCCACAACATATGAAGTATATAGATGGAGAAAAAATTACAAGTAAATGCAGCAAC 1679
Qy 2567 GCAAGTGCTCTAGCCCGCGCTACACTTTCTTC 2600
Db 1680 TGAGTTACCATTAACCTCAAGATACTTCTTATTC 1713

RESULT 4

US-11-216-545-652
; Sequence 652, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: Mc Laird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Soybeans.
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 652
; LENGTH: 2673
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (588)..(588)
; OTHER INFORMATION: n is a, c, g, or t
US-11-216-545-652

Query Match 18.4%; Score 528.6; DB 8; Length 2673;

Best Local Similarity 57.2%; Pred. No. 7.7e-115;

Matches 978; Conservative 0; Mismatches 729; Indels 3; Gaps 1;

Qy 911 ACAAGATATGTAATACCTACGGAACCCACCAAGGCGATAAAGTGGCTTAGGATAC 970
Db 822 AGAGGAATATGCTCACAAGTATGGCCCTACAACTGGTGCACAAAATCCGTCCTTGGTATAC 881
Qy 971 CGATCTTTGGGCGAAGTAGAATGACTATACCACTATGGCGAAGAACTTAATTTGG 1030
Db 882 TGACTTTTGGTGAATTCGAAAGATTTTGGCTCTATGGTGTGATGAATGTGTTTGG 941
Qy 1031 CGCGGGTAAACTATCGTGGGGTATGGCTCAGAGCAATAG---CCCTGATGAAACAC 1087
Db 942 AGTGGAAGATTTTAAAGATGGAATGGTTCAGTCAATGGGGATCCA CTGCAATCTC 1001
Qy 1088 CCTAGATTTTATCATCACTAAACGCGATGATATCGACTACACCGGGATTTTACAAGCCGA 1147
Db 1002 CTTGGATACCTTTATAACAAATGCTGTGATTTCCGCTCGTGGTGGATCATCAAGCGGA 1061
Qy 1148 CATTGGATTTAAAAACGGCAAAATCCATGCGATGGCAAGCAGGAAACAGGACATGCA 1207
Db 1062 TATAGGATTTAAAGATGGTCTCATTTGTTTCAATTTGGAAGAGCAGGAAATCCAGACATCAT 1121
Qy 1208 AGATGGCTAAGCCCTCATATGTCGTTGGTGTGGGCACAGAACCACTAGCAGGGGAAGG 1267
Db 1122 GGAATGATATTTTAAATGATCATTTGGGGCTAATCTGAAGTTATTTGCTGGGAAGG 1181
Qy 1268 TATGATTTATACCGCTGGGGGAATTCGATTCACACCCACTTCTCTCCCAACAATTT 1327

Db 1182 GTTGATTGTAACAGCAGGGGCTATAGATTGTGTCAGTGCATTATATATATGCCCCCAATTAGT 1241
Qy 1328 CCTTACCGCTCTAGCCAATGGCGTTTACAACCATGTTTGGAGGGGCGACAGGTCTCTGTAGA 1387
Db 1242 AGATGAGGGCCATATCAAGCGGCATCAACAATTTAGTGGGAGGTGGAACCGGACCGACTGC 1301
Qy 1388 TGGCAGGAATGGACTACTATCACTCCGGGCAAAATGGAATTCACCGCATGTTGGCGGC 1447
Db 1302 TGGAAACGCTGCGCAACCTTTGTACACGACCATCTCAGATGAAACTTAATGCTGCAATC 1361
Qy 1448 AGCAGAAGAGTAGTATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTTACAAAAA 1507
Db 1362 AACTGATGACCTGCTCTAAACTTTTGGTTTCACTGGCAAGGAAGTAGTTTCCAGCCCTGA 1421
Qy 1508 ACAACTTTGAGAACCAAGTAGAAGCGGGCGCGATTTGGTTTAAATTTGATGAAAGCTGGGG 1567
Db 1422 TGAGCTCCCTCGAGCGGATTCGGCTCGAGCAATGGGACTGAAGCTGCATGAGGACTGGG 1481
Qy 1568 CACAACACCAAGTCCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGCGAAGT 1627
Db 1482 AAGTACTCCCGCTGCAATAGACAGTTGCTTTGACTGTTGCTGATCAATACGATATCCAGAT 1541
Qy 1628 TTGTATCCACACCGCATACAGTCAATGAGGCGAGTTATGTAGATGACACCCCTAAATGCAAT 1687
Db 1542 TAAACATACATACCGACACCTTAATGAGCTGATTTGTGAAACATAGCATTCGAGCAT 1601
Qy 1688 GAAACGGCGCGCATCCATCGCTACCAATGAGGAGCGGGTGGAGGACACTCACCTCGA 1747
Db 1602 TAAAGGAAGAACTATTTCATACCTTACACAGTGAAGGTGAGGTGGTGGTTCATGCTCCAGA 1661
Qy 1748 TGTTATACCATGGCAGGAGCTCAATATTTACCTCTCCACACCCCGCATTTTCC 1807
Db 1662 TATCATCAAGATATGTGGTATGAAGAAATGTTCTGCCCTCATCAACAAACCCCAACCGCCC 1721
Qy 1808 CTATACCATTAATAGCTTGCAGAACACTTAGACATGCTCATGACATGCCACACCTTGA 1867
Db 1722 TTTAACTCAATATATAGATGAGCATCTTGACATGTTGATGGTCTGCCCATCATCTGAA 1781
Qy 1868 CAAACGCTATCCGCGAGGATTTACAATTTTCTCAAAGCCGTATCCGCCCGCGCTCTATCGC 1927
Db 1782 TAGGGAATTCAGAAAGACTTAGCTTTTGCATGTTCAAGATAAGAGAGAAAGGATTCG 1841
Qy 1928 GGCTGAAGATGTCTCCATGATATCGGTGTGATCGCGATGACAAAGTTCGGAATTCGGAAGC 1987
Db 1842 TGCCGAAGATATTTTGCATGATTTTGGGCAATTAGCATCATATCTTCTGATTTCTCAGGC 1901
Qy 1988 AATGGGCGTGCAGCGAAGTATTCCTCGAATTTGGCAGACTCGGATGAAGATTAATAA 2047
Db 1902 TATGGGTGAGTTGGAGAGGTGATGAAGCAGAACTTTGGCAAACTGCCAATAAGATGAAGT 1961
Qy 2048 AGAATTTGTTAAGCTTCTCGAAGATGGCAAGATAAACGATAATTTCCGCAATTAAGCGCTA 2107
Db 1962 ACAACGAGACCACTCCAGCTGGTGAATCCGACATGACAACTTCGGATCAACAGATA 2021
Qy 2108 CATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGTGAAGCAGTATATCGGCTC 2167
Db 2022 CATTGCAAAATACACTATAAATCCAGCTATAGCGAATGGTTTTTTCACATATATGTTGTT 2081
Qy 2168 TGTGAAGAGGGCAAGATCGCGGACTTGGTGTGTGGAATCTCTGCTTTTGGCGGTAAA 2227
Db 2082 GGTGGAGGTGGGAAAGTTAGCTGATCTTTGTGATGTGGAAGCCATCTTTTTTTGGGCGAAA 2141
Qy 2228 ACCCAAAATCGTATCAAGCGGTATGTTGCTCTCTGAAATGGGCGATTTCTAACGC 2287
Db 2142 ACCAGAAATGTTGATCAAGGTGGGGTGGTTCATGGGCTGATATGGGTGACCCCAATGC 2201
Qy 2288 GTCTGTGCCCACTCCCAACCGGTTTATTAACGCGGAAATGTTTGGGCATCACCGCAAGGC 2347
Db 2202 AAGCATCCCCTCTCTGAACCGGTGAAGTGAAGGCTATGTTTGGAACTACTAGCAAGGC 2261
Qy 2348 GAAATTTGACACCGCATCATCTTTTGTTCGAAAGTTCGCTATGAAATGGCGGTGAAGA 2407
Db 2262 TGGTGGTCTTTATCCATAGCTTTTGTGAGCAAGGCGCTGTAGACCAGAGATGATATGC 2321

Qy	2408	AAAGCTGGGCTTTAGAGGCCAAGTTCCTACCGGTCAAAAATGCGGTAAACATACCAAGAA	2467
Db	2322	TCTATACGGACTGAAACAAGAGGGTGGGAAGCAGTAGGCAATGTGAGGAAGCTCTCACTAACT	2381
Qy	2468	AGACTTTCAGTTTCAACGACAAAAACGGCAAAAATCACCGTCGATCGGAAAAACCTTCGAGGT	2527
Db	2382	AGACATGAAACTTTAATGACTCTCTTCACAAATCACTGTGACCCAGATAACTACACTGT	2441
Qy	2528	CTTTGTAGATGGCAAACTCTGCACCTCTAAACCACCTCGCAAGTGCTCTAGCCACGCG	2587
Db	2442	TACAGCAGATGGCGAGGTCTCTACCAGTTTGTGCAACCACTTTGTGCCCTTCTCGAAA	2501
Qy	2588	CTACACTTTCTCTTAGGCACAATGCCCCCT	2617
Db	2502	TTACTTCTCTCTTTTAAGTACTAGCTCCCAT	2531

RESULT 5

```

US-10-449-902-18794
; Sequence 18794, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18794
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK069176
; DATABASE ENTRY DATE: 2001-12-06
; US-10-449-902-18794

```

Query Match	17.7%	Score	509.6	DB	6	Length	2709	
Best Local Similarity	58.7%	Pred.	No. 2.4e-110					
Matches 1007	Conservative	0	Mismatches	684	Indels	24	Gaps	5
Qy	916	AATATGTAATACCTACGGACCCACCAAGGCGATAAGTCGCGTTAGGAGATACCGATC	975					
Db	849	AATATGCAAGCATGTATGGCGCTACAACCTGGTGACAAGATTAGACTTTGGTGACACTGACC	908					
Qy	976	TTTGGGCGAGAAGTAGAACATGACTATACCACCTATGGCGAAGAACTTTAAATTTGGCGCGG	1035					
Db	909	TTTTTGGCGAGATTGAAAGGCACTATGCCATCTATGTTGATGATGATGTCATATTTGGAGCGG	968					
Qy	1036	GTAAAACTATCCGTTAGGGGTATGGGTGAG--AGCA-ATAGCCCTGATGAAAAACACCTTAG	1092					
Db	969	GAAGAAGTTCTACGTGATGGAAATGGACAGTCGGCAGGGGTATCCAGCCTCAGATTGCCTGG	1028					
Qy	1093	ATTTAGTCATCACTTAACCGCATGATTATCGACTACACGGGATTTACAAAGCCGACATTG	1152					
Db	1029	ACACAGTTGTAAACAATATGCTGTTGTGATTGACTATACAGGAATATACAAAGCTGACATTG	1088					
Qy	1153	GGATTAAAAACGGCAAAATCCATGGCATTTGGCAAGCAGGAAAAAC---AAGGACATGCA--	1207					
Db	1089	GTATAAATGGTGACCTTATCGTTGCTATTGGGAAGGCTGGNAACCCCTGATGTATGGACA	1148					
Qy	1208	-AGATGCGGTAAGCCCTCATATGCTCGGTGTGGGTCAGAGCACTAGCAGGGGAAG	1266					
Db	1149	TGGATGCGTTAATGAAGAGATGATTGTGGGGGTTAATCTCGAATTTATGCTGCTGAAG	1208					

Qy	1267	GTATGATTATTACCGCTGGGGGAATCGATTTCACACCCACTTCCTTTCTCCACAACAAT	1326
Db	1209	GCATGATTGTTACTGCTGGTGGAAATGATTGCCATGTTTCACTTCATATGCCTCAGTTGG	1268
Qy	1327	TCCCTACCGCTCTAGCCAAATGSGCTTACAAACCATGTTTGGAGGCGGCACAGGTCCCTGAG	1386
Db	1269	CAGAAGGCAATTGCAAGTGGCATCACAACTTGGTGGGAGGTGGAACTGGACACGAC	1328
Qy	1387	ATGGCACGAATGGCACTACTATCACTCCGGGCAAAATGGAACTTGCACCGCATGTTGGCGG	1446
Db	1329	ATGGAACTTGTGCCACAACCTTGCACCCCTTCCACATCTCATAGAACTAATGCTACAGT	1388
Qy	1447	CAGCAGAAGAGTATCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAA	1506
Db	1389	CCACTGATGAATTACCAATCAATATGGGATTACAGGCAAGGGGAAATACTACAAAACCTG	1448
Qy	1507	AACNACTTGTAGAACAAAGTAGACGGGCGGATTTGTTTTAAATTCGATGAGACTGGG	1566
Db	1449	ATGGATTGGCTGAGATCATTTAAGCGAGGACAATGGGCTTGAAGTTCGATGAAGATTGGG	1508
Qy	1567	GCACAACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAG	1626
Db	1509	GAAGCACCCCGCTGGCATAGATAACTGTGTTACTGTGCGAGAAGCTTTGATATCCAGG	1568
Qy	1627	TTTTGTATCCACACCGATPACAGTCAATGAGCGAGGTTATGTAGATGACACCGCTAAATGCAA	1686
Db	1569	TAAATATCCACACAGACACCTTAAATGAGTCAGGCTGTGTGGAGCATACAATTTGCCGAT	1628
Qy	1687	TGAACGGGCGGCATCCATGCCCTACACATTTGAGGAGCGGTGGAGGACACTCACTG	1746
Db	1629	TTAAGATAGGACAATACATACATATCATAGTGAAGGAGCAGGTGGTGCATGCTCCAG	1688
Qy	1747	ATGTTATCACCATGGCAGGCGAGCTCAATATTCTACCCCTCTCCACCACCCCACTATTTC	1806
Db	1689	ACATTATCAAGTATGTGGAGTGAANAATGTGTTACCTCTTCAACAAATCCAACTCGGC	1748
Qy	1807	CCTATACCAATTAATACGTTTGCAGAACACTTAGACATGCTCATGACATGCCACCACTAG	1866
Db	1749	CAITTTACTTTTGAACACTGTAGATGAGCACCTTGATATGCTGATGCTGCCATCACCTCG	1808
Qy	1867	ACAAACGCATCCGGGAGGATTTACAATTTTCTCMAAGCGGTATCCGCCCGCGGCTCTATCG	1926
Db	1809	ATAGAAATATCCCAGAGAGATGTAGCATTTGCCAGAGTCTAGAAATTCGAGCTGAACAATAG	1868
Qy	1927	CGGCTGAAGATGTGCTCCATGATATGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAG	1986
Db	1869	CTGCTTAGGATATCTTGCAATGACATGGAGCAATTCAGTATCATATATCCGATTCGCAGG	1928
Qy	1987	CAATGGGGCTGCAGGCGAAAGTATTCCTCGAACTTGGCAGACTCGCGATAAGAAATAAAA	2046
Db	1929	CTATGGGGCGCAATTGGAGAGGTGATTACTCGGACATGCGCAAACTGCAAAATAAGATGAAGA	1988
Qy	2047	AAGAAATTTGTAAGCTTTCCT-----GAAAGATGGCAAAAGATAACGATAATT	2091
Db	1989	GACAGAGAGGTAGATTACCTATATATCCAGTTTCCCGATGCTGTCAGAGGACAATGACAATT	2048
Qy	2092	TCCGATTATAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGCGTGA	2151
Db	2049	TCCGGATAGAAGAGATACATAGCTAAATACACCATAAATCTCTGCTATATAGTGAATGGCTTTT	2108
Qy	2152	GCAGATATATCGGCTCTGTGGAAAGAGGGCAAGATCGCCGACTTCGTGGTGTGGAAATCCTG	2211
Db	2109	CTGATTTTGTGTGCTCTGTTGAGGTGGGAAAATTAGCTGACCTTGTTATTTTGGAAACCTT	2168
Qy	2212	CCTTTTTGTGCGTAAAAACCAAATCGTGTGATCAAAGCGGTATCGGTGGTCTTCTCTGAAA	2271
Db	2169	CTTTCTTTTGGAGCAAAACCAAGAAATGGTTATAAAGGGTGGTGTCTATTGCAATGCGCTAATA	2228
Qy	2272	TGGCGGATTTCTAACGGCTGTGTGCCACTCCCCAAACCGGTTTATTAACGCGAAATGTTTG	2331
Db	2229	TGGGTGATCCAAATGTGATGCAATTCACACCTGAAACCGGTTATGATGCGACTTATGTTTG	2288

Qy 2332 GGCATCAAGCGAAGCGAAATTTGACACGAGCATCACTTTTGTCTTCCAAAGTCGCTATG 2391
Db 2289 GTGCATTTGGAGGGCTGGAAGTGCCTAATAGCAATTTGTGAGCAAGGCTGCTAAG 2348
Qy 2392 AAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAACTGCC 2451
Db 2349 AAGCTGGGTTGCAGTGCAGTACAAAGCTGGGAAAGAGGTTGGAAGCTGTAGTCCGGTTC 2408
Qy 2452 GTAACATCACCAAGAAAGACTTCAAGTTCAACGACAAACGGCAAAAAATCACCGTCGATC 2511
Db 2409 GTGGTTTGACAAAGCTGAATATGAATCAAACTTAAACGATGCACTTCCGAAAAATTTGATGTCGATC 2468
Qy 2512 CGAAACCTTCGAGTCTTCTGATGTCGAACACTCTGCACTCTTAAACCCACCTCGCAAG 2571
Db 2469 CTGAACCTCACCGTTACTGCTGATGAGAGGTTTTCAGATGTCAACCAACCCACAG 2528
Qy 2572 TGCCTCTAGCCAGCGGTACACCTTTCTTCTTAGGCA 2606
Db 2529 TACCACTATCTCGGAATTAATCTTCTCTTTAGACA 2563

RESULT 6

US-11-218-305-14860
; Sequence 14860; Application US/11218305
; Publication No. US20060141495A1

GENERAL INFORMATION:

; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: Tao, Nengbing
; APPLICANT: Mc Laird, Paul L.

; APPLICANT: Wu, Kunsheng

; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn.

; FILE REFERENCE: 38-21 (536660)B

; CURRENT APPLICATION NUMBER: US/11/218,305

; CURRENT FILING DATE: 2005-09-01

; PRIOR APPLICATION NUMBER: US 60/606,880

; PRIOR FILING DATE: 2004-09-01

; NUMBER OF SEQ ID NOS: 25043

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 14860

; LENGTH: 3120

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (16)..(16)

; OTHER INFORMATION: n is a, c, g, or t

US-11-218-305-14860

Query Match 17.2%; Score 495.4; DB 9; Length 3120;
Best Local Similarity 57.6%; Pred. No. 5.8e-107;
Matches 982; Conservative 0; Mismatches 706; Indels 18; Gaps 3;

Qy 916 AATATGTAATACCTACGACCCCAAAAGCGGTAAGTTCGTTAGGATACCGATC 975
Db 1083 AGTATGCGAGCAATTATGAGACCTACCAACCGGTGATAAGATTAGGCTTGGTGATACCAATC 1142
Qy 976 TTTGGCAGAGTAGACATGACTATACCACTATGCGGAGAACTTAAATTTGGCGGG 1035
Db 1143 TTTATGCTGAGATTGAAGAGACTTCGCCCTTTTATGTTGATGAGTGATATTTGGCGGTG 1202
Qy 1036 GTAAATATCTCCGTGAGGGTATGGTTC-AGAGCAATAG--CCCTGATGAAACACCCCTAG 1092
Db 1203 GAAAGATTCTCGGTGATGGATGGGACAAAGCTTCAGGGTACCCAGATCTTCTTCGCTAG 1262
Qy 1093 ATTATGATCATCAACGCGGATGATTATCGACTACACCGGGATTTACAAAGCCGACATTG 1152
Db 1263 ATACAGTTATAACAAATGCTGTTGTGATTGATTATACCGGAATATACAAGGCTGATATCG 1322
Qy 1153 GGAATTAACCGCAAAATCCATGGCTTGGCAGGCGAGGAAACAGGACATGCAAGATG 1212
Db 1323 GTATAAAGGTGAGCTTTATAGTTGCTATTGGAAAGGCTGGAAACCCCTGATGTCATGATG 1382

Qy 1213 GGGTAAGCCCTCATATGGTCTGGGTGTGGGACAGACACTAGCAGGGGAGGTATGA 1272
Db 1383 GTGTCCATAAACACATGATTTGGGGTCAATACTGAAGTTATTTGCACTGAAGGCATGA 1442
Qy 1273 TTATTACCCCTGGGGAATCGAATTCACACCCACCTTCTTCTTCCCAACAATTTCCCTA 1332
Db 1443 TTGTCACTGCTGGTGGCAGTAGATTGCCATGTCCACTTTATATGTCTCTCAGTTGGCAGAAG 1502
Qy 1333 CGGCTCTAGCCAAATGGCGTTACAAACATGTTTGGAGGCGGACAGGTCTCTGTAGATGGCA 1392
Db 1503 AGGCAATTGCAAGCGGCATCACGACATTTGGTGGTGGTGGAAACCGGACAGCAGCATGAA 1562
Qy 1393 CGAATGCGACTACTATCACTCCGGGCAATGGAACCTTGCACCCCATGTTTGGGCGCAGAG 1452
Db 1563 CTTGTGCCCAAACTTGCACCTCCAGCAACCATCTCAATTGAAATTAATTTGTCAGCTCCA 1622
Qy 1453 AAGAGTATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAAC 1512
Db 1623 ATCAATTGCCAATTAACTGGGANTTCAAGGCAAGGGAATACTTCAAAACCTGAAGGAT 1682
Qy 1513 TTGTAGAACAGTAGAAGCGGCGGATTTGTTTAAATTTGCATGAAGACTCGGGGCAAA 1572
Db 1683 TGGCTGAATCATTAAGCTGGAGCAATGGGTTTGAAGCTGCATGAAGATTGGGGAACATA 1742
Qy 1573 CACCAAGTGGATCGATCACTCTTGGAGCGTGGCAGATGAATAGATGTGCAAGTTTGTGA 1632
Db 1743 CCCCATCCGCGATAGATTAATTTGTTATCTGTTGCGAGAAGATTTTGTATATTTCAGGTCAATA 1802
Qy 1633 TCCACACCGATACAGTCAATGAGGAGGATTATGTAGATGACACCCCTAAATGCAATGAACG 1692
Db 1803 TACACACAGATACCTTAAATGAATCAGGCTGCGTGGGAGCATACATCGCAGCTTTTAAAG 1862
Qy 1693 GCGCGCCATCCATGCTACACACATTTGAGGAGCGGCTGGAGGACACTCACCTGATGTTGA 1752
Db 1863 GTAGAGCCATACATACATATCACAGTGAAGTGTCTGTGGCGGTCTGCTCCAGACATCA 1922
Qy 1753 TCACCATGGCAGCGAGCTCAATATTCTACCTCTCTCCACACCCCACTATTTCCTTATA 1812
Db 1923 TCAAAGTTTGTGGGGTAAATAATGTTGCTCCCTCTTCAACAAATCTTACCGGCCATTTA 1982
Qy 1813 CCATTAAATACGGTTGCGAACAACACTTAGACATGCTCATGACATGCCACCACTAGACAAAC 1872
Db 1983 CTTCGAACACTGTAGACGAGCACCTTGATATGCTGATGTTTGTACCACTTGTATAAAA 2042
Qy 1873 GCATCCGCGAGGATTACAAATTTCTCAAAGCGGTATCGCCCGGCTCTATCGCGCTG 1932
Db 2043 ACATTCAGAGATGTAGCATTTGCGAGTCTAGAAATTCAGCTGAAACAATTTGCTGCTG 2102
Qy 1933 AAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGG 1992
Db 2103 AGGACATATTGCATGACATGGAGCCATAAGTATTATATCATCTGATTCACAGGCCATGG 2162
Qy 1993 GCGGTGAGGCGAAGTGTCTCTCGAACTTGGCAGACTGGCGGATAGAAATAAAAAAGAT 2052
Db 2163 GCGCGTTTGGAGAGGTGTATAACGAACTGCGCAAACTGCAAAACAGATGAAGGTACAAA 2222
Qy 2053 TTGTTAAGCTTCTCGAAGATGCAAGATAA-----CGATAATTTCCGCA 2097
Db 2223 GAGGTGTTTACCTGGATCTGTTGAGCTATGCGCCCGGACAGTGAACCTCCGTA 2282
Qy 2098 TTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCCAAGCGGTGAGCGAGT 2157
Db 2283 TAAGAGATACATAGCAAAATACAAATATCCAGCAATAGTGAACGGGTTTTCAGACT 2342
Qy 2158 ATATCGGCTCTGTGGAAGAGGCAAGATCGCGCACTTGGTGTGTGGAAATCTTGCCTTTT 2217
Db 2343 TTGTTGGTCTCTGTTGAGGTGGAAAAATTAGCTGACCTTGTCTTTTGGAAACCACTTTCT 2402
Qy 2218 TTGGGTAAACCCCAAAATCGTGAAGGCGGTATGGTGTCTTCTCTGAAATGGGG 2277
Db 2403 TCGGCGCTAAACCCAGAACTGGTTGTAAGGGGAGGTGCAATTCATGGGGTAAACATGGGTG 2462
Qy 2278 ATTCTAAGCGCTCTGTGCCCACTCTCCCAACCGGTTTATTACCGGAAATGTTTGGGCATC 2337

[illegible]

RESULT 7

US-10-953-349-33824
; Sequence 33824, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33824
; LENGTH: 2912
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33824

Query Match	17.1%; Score 493.8; DB 6; Length 2912;
Best Local Similarity	.57.5%; Pred. No. 1.4e-106;
Matches 981; Conservative	0; Mismatches 707; Indels 18; Gaps 3;
Qy	916 AATATGTAATACTACGGACCCACCAAGGCGATAAGTCGCTTAGGAGATACCGATC 975
Db	
Db	1067 AGTATCGAGCATTTATGSAACCTACCAACCGGTGATAGATTAGGCTTGGTGATACCAATC 1126
Qy	976 TTTGGGCAGAAGTAGAACATGACATATACCACCTATGCGAAGAACTTAAATTTGGCGCGG 1035
Db	
Db	1127 TTTATGCTGAGATTGAAAGACATTCGCTTTTATGGTGATGATGTCATATTGGCGGTG 1186
Qy	1036 GTAAACATATCCGTGAGGGTATGGGTC-AGAGCAATAG--CCCTGATGAAAAACACCCCTAG 1092
Db	
Db	1187 GAAAAGTTCTGGGTGATGGCATGGACAAGCTTCAGGGTACCCAGAATCTTTCTGCCTAG 1246
Qy	1093 ATTTAGTCATCACTAAACGGATGATTATCGACTACACGGGATTTACAAAGCCGACATTG 1152
Db	
Db	1247 ATACAGTTATAACAAATGCTGTTGTGATTGATTATACCGGAATATACAAGGCTGATATCG 1306
Qy	1153 GGATTAACAAACGGCAAAATCCATGGCATTCGCAAGGCGAGGAACAAGACATGCAAGATG 1212
Db	
Db	1307 GTATAAAGGTGGACTTATATAGTTCTATTTGAAAGGCTGGAAACCCCTGATGTCATGGATG 1366
Qy	1213 GCGTAAGCCCTCATATGCTGTGGGCTGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGA 1272
Db	
Db	1367 GTGTCCATAACACATGATTTGTTGGGGTCAATACTGAAGTTATTGCATCTGAAGGCATGA 1426
Qy	1273 TTAATTACCGCTGGGGGAATCGATTCAACACCCCACTTCCTTTTCTCCACAACAAATTCCTTA 1332

1427	Db	TTGTCACTGCTGCTGGGATAGATTGCGCATGTCCACTTTTATATGTCTCTCAGTTGGCGAGAAG	1486
1333	Qy	CCGCTCTAGCCAATGGCGTTTACAACCATGTTTGGAGGCGGCACAGGTCTCTGTAGATGGCA	1392
1487	Db	AGGCNATTGCAAGCGGCATCAGACATTTGGTGGTGGTGGAAACGGACCCAGCACATGGAA	1546
1393	Qy	CGAATGCGACTATATCATCTCGGGGCAAAATGGAACTTTGCAACCGCATGTTGGCGCGCAGCAG	1452
1547	Db	CTTGTGCGCAACAATTGCACTCTCAGACGACCATCTCAATTTGAAATTTAAATGTTTGCAGTCCCACTG	1606
1453	Qy	AAGAGTATTCTATGAATGTGGGCTTTTGGCGAAAGGCAATAGCTCTAGCCAAAACCAAC	1512
1607	Db	ATCAATTTGCCAATTAACATGGGAATTCAGGCAAGGGAATATCTTCAAAACCTTGAAGGAT	1666
1513	Qy	TTGTAGAAACAAGTAGAAGCGGCGCGCATTTGTTTAAATTTGCAATGAAGACTGGGGGCACAA	1572
1667	Db	TGGCTGAATCATTAAGCTGGAGCAATGGGTTTGAAGCTGCATGAAGATTGGGGAACATA	1726
1573	Qy	CACCAAGTGCATTCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAAGTTTGTGA	1632
1727	Db	CCCCATCCGCGATAGATAAATTTGTTTATCTGTTGCGAAGAATTTTGTATTTTCAGGTCAATA	1786
1633	Qy	TCCACACCGATCAGTCNAATGAGCGAGGTTATGCTAGATGACACCCCTAAATGCAATGAACG	1692
1787	Db	TACACACAGATACCTTAAATGAATCAGGCTGGCTGGAGCATACAATCGCAGCTTTTAAAG	1846
1693	Qy	GGGCGGCATCTCATGCTACCAATTTAGGGGAGCGGTGGAGGACACTCACCTCATGTTTA	1752
1847	Db	GTAGAGCCATACATACATATCAAGTGNAGGTGCTGGTGGCGTCACTGCTCCAGACATCA	1906
1753	Qy	TCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCACTATTTCCCTATTA	1812
1907	Db	TCAAAAGTTGTGGGTAAAAAAATGTGTGGCTCTTTCAACAAAATCTTACCCGGCCATTTA	1966
1813	Qy	CCATTANTAAGTTTGCAGACACTTAGACATGCTCATGATGCCACCACTAGACAAAC	1872
1967	Db	CTTCGAACACTGTAGACGAGCACTTGATATGCTGATGGTTGTGCACCACTTGATATAAA	2026
1873	Qy	GCATCCGCGAGGATTTACAATTTTCTCAAAAGCGATATCCGCCCGGCTCTATCGCGGCTG	1932
2027	Db	ACATTTCCAGAAATGTAGCATTTGCCGAGTCTAGAATTCGAGCTGAAACAATTTGCTGCTG	2086
1933	Qy	AAGATGTCTCATGATATGGGTGTGATCCGATAGCAAGCTCGBATTCGCAAGCAATGG	1992
2087	Db	AGGACATATTGTCATGATCGGAGCCATAGTATTTATCATCTGATTCACAGGCCCATGG	2146
1993	Qy	GGCGTGCAGCGGAAGTATTCCTCGAACTTGGCAGACTCGCGATAGAAATAAAAAAGAT	2052
2147	Db	GGGCGCTTGGAGAGGTGATTAACACGAACATGGCAAACTGCAAAACAAGATGAAGGTACAAA	2206
2053	Qy	TTGTGAGCTTCTGTAAGATGGCAAAAGATAA-----CGATAAATTTCCGCA	2097
2207	Db	GAGGTAGTTTACCTTGTGATCTGGTGACGCTAATGCGGCCCGGACAGGTGACAACTTCGTA	2266
2098	Qy	TTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGAACCAAGCGGTGAGCGAGT	2157
2267	Db	TAAGAAGATACATAGCAAAAAATACAAATAAATCCAGCAATAGTGAATGGGTTTTCAGACT	2326
2158	Qy	ATATCGGCTCTGTGGAAGAGGCAAGATCGCCCACTTGGTGGTGTGGAATCTTCGCCCTTTT	2217
2327	Db	TTGTGTGTTCTGTGTGAGGTGGGAAAAATTAGCTGACCTTGTCTTTTGGAAACCACTCTTCT	2386
2218	Qy	TTGGCGTAAACCCAAAATCGTGAATCAAAAGCGGTATGGTGGTCTTCTCTGAAATGGGCG	2277
2387	Db	TCGGCGCTAAACGAGAACTGGTTGTAAAGGAGGTGCAATTTGATGGGCTTAACATGGGTG	2446
2278	Qy	ATTCTAAACGGCTGTGCGCCACTCCCCAACCCGGTTTATTATCCGCGAAATGTTTGGGCATC	2337
2447	Db	ATCCCAATGCTAGCATTTCCAACACCTTGAACTGTTGTGATGCGACCTATGTTTGGTGCAT	2506
2338	Qy	ACGCAAGGCGAAATTTGACACCAAGCATCACTTTTGTGTTTCCAAAGTCGCTATGAAATG	2397
2507	Db	TTGAAAGGCTGGAAGTTCCAAATTCAAATGCAATTTGTGAGCAAGGCTGTAAAGAAGCTG	2566

Qy	2398	GCCTGAAAGAAAAGCTGGCTTTAGAGCGCCCAAGTTCTTACCGGTCAAAAACACTGCCGTAAACA	2457
Db	2567	GTGTGCGCAACGGAGTTACAGACTAGAAAGAGGGTGGAAAGCTGTAGGCCGTGTTTCGAGGCC	2626
Qy	2458	TCCACCAAGAAGACTTCAAGTTCAACGACAAAACGCGAAAATCAACCGTTCGATCCGAAAAA	2517
Db	2627	TGACAAAGCTTGACATGAAGCTCAACGACGCCCTCCCGAAAAATAGAGGTGCGACCCGGAGA	2686
Qy	2518	CCTTTCGAGGTCTTTGTGTAGATGGCAAACTCTGACACTCTAAACCCACCTTCGCAAGTGCCTC	2577
Db	2687	CCTACACAGTGAACCGCTGATGGAGAGGTTTTTGACATGCCAACCAACCCACCACTGCGCAC	2746
Qy	2578	TAGCCCAAGCGCTACACTTCTCTTCAG	2603
Db	2747	TGTCTCGGAATTACTTCTCTGTTCTAG	2772

RESIST. 8

```

RES001.8
US-11-292-431-36
; Sequence 36, Application US/11292431
; Publication No. US20060121061A1
;
GENERAL INFORMATION:
;
; APPLICANT: COLE, Garry T.
; APPLICANT: CHEN, Xia
; APPLICANT: SESHAN, Kalpathi R.
; APPLICANT: HUNG, Chiung-Yu
; APPLICANT: XUE, Jianmin
; APPLICANT: YU, Jieh-Juen
;
TITLE OF INVENTION: Attenuated Vaccine Useful for
; Immunizations Against Coccidioides spp. Infections
;
FILE REFERENCE: 529522006000
;
CURRENT APPLICATION NUMBER: US/11/292,431
;
CURRENT FILING DATE: 2005-12-02
;
PRIOR APPLICATION NUMBER: US 60/633,399
;
PRIOR FILING DATE: 2004-12-03
;
NUMBER OF SEQ ID NOS: 37
;
SOFTWARE: FastSeq for Windows Version 4.0
;
SEQ ID NO 36.
;
LENGTH: 5588
;
TYPE: DNA
;
ORGANISM: Coccidioides posadasii
;
US-11-292-431-36

```

Query Match	16.8%;	Score 483.4;	DB 8;	Length 5588;
Best Local Similarity	60.3%;	Pred. No. 5.1e-104;		
Matches 832;	Conservative 0;	Mismatches 541;	Indels 6;	Gaps 2;
QY	902	AATGAAAAAACAAGATATGTAAATACCTTACGACCCACCAAGAGCGGATAAAGTGC	961	
Db	2636	AATGACGCGGAAGCATATGCTCGAATGTTTGGTCTTACCACCTGGAGATGTAGTCAAGCT	2695	
QY	962	AGAGATACCGATCTTTGGGAGAGTAGAACAATGACTATACCACTTATGGCGAAGAACT	1021	
Db	2696	AGAAACCAAGATTTTGTGGATTAAGTCGAAAAGGACCTGACCTACTATGTTGACGAATG	2755	
QY	1022	TAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTCA---GAGCAATAGCCCTGA	1078	
Db	2756	TTCAATTCGGTGTGGCAAGACCATAAGAGACGGGATGGGGCAAGCTACAGGAAGGCATTC	2815	
QY	1079	TGAAAAACACCTAGATTTTAGTTCATCACTAAGCGGATGATTTATCGACTACACCGGATTTA	1138	
Db	2816	CGTGGATGTCTCGATACAGTCTCTGTGTGAACGGCTAAATTTGTGATTTGGACGGGTATTTA	2875	
QY	1139	CRAAGCCGACATTTGGGATTAATAAACGSCAAAATCCATGGCATTTGGCAAGCAGGAACAA	1198	
Db	2876	CAAGCTGATTTTGGACTAAAGAATGATTTGATCTCGCGAATCGGCAAAAGCTGGAACCC	2935	
QY	1199	GGACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGTGGGTGTGGGCACAGAAGCACTAGC	1258	
Db	2936	AGACATGATGGATGGTGTCAACCCCAACATGATAGTTGGCTCTTTCGACAGATGTTATCGC	2995	
QY	1259	AGGGGAAGGTATGATTTATTACCGCTCGGGGAATCGAATTTCAACACCCACATTCCTTTCTCC	1318	

RESULT 9
US-11-348-413-11113/c
; Sequence 11113, Application US/11348413
; Publication No. US20060160121A1
; GENERAL INFORMATION:


```
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 60/615,573
; PRIOR FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 11113
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(315)
; OTHER INFORMATION: WANO1UFWB; ORF: 0000000011100; Cluster contains WANO1UOT1
; OTHER INFORMATION: egad_360118_2391:EGAD orf prediction:ATCC12228.NC_004461
; OTHER INFORMATION: .1
US-11-348-413-11113

Query Match      3.3%; Score 94.4; DB 9; Length 315;
Best Local Similarity 58.9%; Pred. No. 1.3e-12;
Matches 188; Conservative 2; Mismatches 123; Indels 6; Gaps 1;

Qy 1114 TGATTATCGACTACACCGGGATTACAAAGCCGACATTTGGGATTAATAAAGCGCAAAATCC 1173
Db 314 TAATTATTGATTATGACAAGATTCTTAAAGCAGATATCGGAATTAATAAATGTTATATT 255

Qy 1174 ATGCCATTTGGCAAGCGAGAAACAGGCATGCAAGATGGCGTAAGCCCTCATATGGTCG 1233
Db 254 TTAAGATCGGTAAAGCTTGGAAACCCAGATATAATGGATAACGT-----TGACATCATCA 201

Qy 1234 TGGGTGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTATACGCTCGGGGAATCG 1293
Db 200 TTGGTGCACACACTGATATTAATGCTGCTGAAGGTAAATTTGTTACTGCCGGCGGTATCG 141

Qy 1294 ATTCACACACCCACTTCCTTTCTCCACAACTTCCCTACCGCTCTAGCCCAATGGCGTTA 1353
Db 140 ATACACAGTGCATTTATCAATCTGACACAGCTGAGTTGCACTTGAGAGTGGTATTA 81

Qy 1354 CAACCATGTTTGGAGCGGCGACAGGTCTCTGATAGTGGCAAGATGGCGACTACTATCACTC 1413
Db 80 CAACGCATATCGTGGAGGAACCTGCTGCTTCTGAAGGTGCTTAAAGCGACTACTGTAAACAC 21

Qy 1414 CGGCAATGGAACCTTCCA 1432
Db 20 CAGGACCTTGGCATATTTCA 2

RESULT 10
US-11-348-413-12794
; Sequence 12794, Application US/11348413
; Publication No. US20060160121A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
```

```
; PRIOR APPLICATION NUMBER: US 60/615,573
; PRIOR FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 12794
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(321)
; OTHER INFORMATION: WANO1UOV8; ORF: 00000000011100; Cluster contains WANO1OSHMH
; OTHER INFORMATION: SE1861:urea:urease gamma subunit:ATCC12228.NC_004461.1
US-11-348-413-12794

Query Match      3.2%; Score 92.6; DB 9; Length 321;
Best Local Similarity 58.7%; Pred. No. 3.5e-12;
Matches 176; Conservative 1; Mismatches 120; Indels 3; Gaps 1;

Qy 207 TGAAGCTCACACCCAAAGAGCAAGAAAAGTTCTTGTATATATATGCGGGGGAAGTGGCTA 266
Db 20 TGCACCTTTACACACAGCTGAACAAAGACAAATTGATGATGTTGTAGCTGCTGAGGTTGCAC 79

Qy 267 GAAAGCGCAAAAGCACAGAGGCTTTAAAGCTCAACCAACCGGAAGCCATTGCTTACATTAGTG 326
Db 80 GTGCTAGAAAAGCAAGAGACTTAAACTTAATCATCTGAAGCACTTGCTTTAATCAGTG 139

Qy 327 CCCATATTATTGACGAAGCGCGCCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAGT 386
Db 140 ATGATATTATTAGAAGCGCGCGTGATGG--TAAACCGTAGCTGAATVATGAGCTATG 196

Qy 387 GCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTAAATATGTTCCCGATC 446
Db 197 GAAAAACAATTTTAAACGAGGAAGATGTCATGGATGGCGTAGCTAAACATGATTACAGAAC 256

Qy 447 TAGGTGTAGAGCCACCTTTCTCTGATGGTACGAAACTTGTAACTGTGAATTGGCCCATCG 506
Db 257 TTGAATTTGAAGCAACTTTTCCAGATGGTACTAAGTTAATAACAGTCCATCACCCCAATCG 316

RESULT 11
US-10-471-571A-3255
; Sequence 3255, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 3255
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3255

Query Match      3.1%; Score 89.8; DB 6; Length 300;
Best Local Similarity 58.3%; Pred. No. 1.6e-11;
Matches 175; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

Qy 207 TGAAGCTCACACCCAAAGAGCAAGAAAAGTTCTTGTATATATATGCGGGGGAAGTGGCTA 266
Db 2 TGCATTTTACACAGAGAGCAAGACAAATTAATGATTGTAGTGGCGGCGGAAGTTGCAC 61

Qy 267 GAAAGCGCAAAAGCACAGAGGCTTTAAAGCTCAACCAACCGGAAGCCATTGCTTACATTAGTG 326
Db 62 GTGCTGCTAAGACACGTTGTTTGAACATAAATCATCTCTGAGGCAATTAGCTTTAATCAGCG 121

Qy 327 CCCATATTATTGACGAAGCGCGCCCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAGT 386
```

Db 122 ATGAATTATTAGAGGTGCACCGATGG--TAAGACCGTTGCAGAGTTAATGAGTTATG 178
Qy 387 GCATGCACTTTTGAAGAAAGATGAATATGCCCGGGTGGGTAATATGTTTCCCGATC 446
Db 179 GTAGACAAATCTTAACAAAGAGATGTCTGGATGGTGTGCAACACATGATTACAGATA 238
Qy 447 TAGGTGTAGAGCCACCTTTCTCATGTGTACGAACTTGTAACTGTGAATGGCCCATCG 506
Db 239 TCGAAATCGAGGCTACGTTCCCGATGGTACTAAGTTAATCAGATACATCACCCCTATTG 298

RESULT 12

US-11-348-413-11115
; Sequence 11115, Application US/11348413
; Publication No. US2006016012A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 60/615,573
; PRIOR FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 11115
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(402)
; OTHER INFORMATION: WANO1UFWI; ORF: 00000000011110; Cluster contains WANO1OSH
; OTHER INFORMATION: :SE1862:ureb beta subunit:ATCC12228:NC_004461.1
US-11-348-413-11115

Query Match 3.0%; Score 85.2; DB 9; Length 402;
Best Local Similarity 52.8%; Pred. No. 2.2e-10;
Matches 201; Conservative 2; Mismatches 175; Indels 3; Gaps 1;
Qy 543 TTGGTTCGATAAAGCATCGAGCTCAATCGCAGCAAGAGTAACCGAAGTTCAGGTTA 602
Db 20 TTGTTAAAAATACTGAAATAGAGTCAATAAATCAATCATCCCGAAACGGTGAATGAAGTGA 79
Qy 603 CTAATGAAGGCGCTAAATCCTTTCATGTGGTAGCCATTTCCACTTCTTTTGAAGCTAACA 662
Db 80 AAAATACRGGGATAGACCTATACAGTAGGTTACAAATTTCCACTTTTTCGAAACAATA 139
Qy 663 AGGCACATAAATTCGATCGTGAAGAAAGCCTATGCAAGCGCTAGATATTCCTCTGCGCA 722
Db 140 ARGCAATTAGATTTGATCGTGAGAAACATATGTAACATTTGGATATTCCTCGCAGGAG 199
Qy 723 ACAGCTACGATTTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCA 782
Db 200 CTGCAGTGAGATTGAACTTGGAGATGAAAAAAGTACAACTTGTGCAATATTTCTGGAC 259
Qy 783 GTAAAAAGCTGATGGCATGACGCGGCTTGTAAT---AACATCGCGATGACGCCATA 839
Db 260 GACGTAAAAATTTATGGATTCGTTGGTGTAGTCGATGCGGATATTGACGAAAGACCGGTAT 319
Qy 840 AACATAAAGCGCTTGACAGGCGAAATCTCACCGATTTATCAAGTAAGGAGACTCCCATG 899
Db 320 TCGTTCGAATGATTCAAATCAAAACGCCCGGTTAAAAACGATGACGGCGAGACAATG 379
Qy 900 AAAATGAAAAAACAAGAAATAT 920
Db 380 CGAATAAAAAAGGTGTAAT 400

RESULT 13

US-10-471-571A-3277
; Sequence 3277, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3277
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-3277

Query Match 2.8%; Score 80.2; DB 6; Length 408;
Best Local Similarity 50.9%; Pred. No. 3.3e-09;
Matches 190; Conservative 0; Mismatches 183; Indels 0; Gaps 0;
Qy 553 TAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAACCGAAGTTCAGGTTACTAATGAAGG 612
Db 30 TACAGAGGTTGAAATTAATAACCATCATCTCTGAACAGTTATCGAAGTTGAAAAATACAGG 89
Qy 613 GCCTAAATCCTTTCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCACCTAAA 672
Db 90 AGATCGACCAATTCAGAGTGGGCTCAATTTTCATTTTATGAAGCAAAATGCGACATTTAGA 149
Qy 673 ATTGCATCGTGAAGAAAGCTATGCGAAAGCCTAGATATTCCTCTGCGCAACACGCTACG 732
Db 150 TTTCGACGTGAAATGGCATATGGAAGAACATTTAGATATTCAGCTGGAGAGCTGTTCG 209
Qy 733 CATTCGGGCGAGACAAACCCGAAAGTGCAGTTGATTCCTCTTTGGTGGCAGTAAAAAAGT 792
Db 210 ATTGAACTGGGATAAAGAAAGATTCATTTAGTATGCTGGCAACGTTAAAT 269
Qy 793 GATTGGCATGAACGGGCTTGTGAATAACATCGCGAGTAACCGCCATAAACAATAAAGCGCT 852
Db 270 TTTTGGTTTTCGTGTATGGTCAATGGTCTTATCGATGAGTCACGCTGTCTATCGCCCAAC 329
Qy 853 TGACAGCGGAATCTCAGGATTTTATCAAGTAAGAGACTCCCATGAAATGAAAAAAC 912
Db 330 TGATGAAATGATGAATATGCGAGGTGTATTTCGAGATAACCGGTGCTGAAAAACGTGAATA 389
Qy 913 AAGATATGTAAA 925
Db 390 AAAGGAGAAAA 402

RESULT 14

US-11-348-413-11114/c
; Sequence 11114, Application US/11348413
; Publication No. US2006016012A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084100 (AM 101724)
; CURRENT APPLICATION NUMBER: US/11/348,413
; CURRENT FILING DATE: 2006-02-07
; PRIOR APPLICATION NUMBER: PCT/US05/035471
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 11/243,445
; PRIOR FILING DATE: 2005-10-05
; PRIOR APPLICATION NUMBER: US 60/615,573

```

; PRIOR FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 11114
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(240)
; OTHER INFORMATION: WANO1UFWE; ORF; 00000000011100; Cluster contains WANO1UO0U
; OTHER INFORMATION: : egad_359698_2389::EGAD orf prediction:ATCC12228:NC_004461
; OTHER INFORMATION: .1
US-11-348-413-11114

```

Query Match	2.3%	Score 66.2;	DB 9;	Length 240;
Best Local Similarity	56.9%;	Pred. No. 5.4e-06;		
Matches 119;	Conservative 1;	Mismatches 89;	Indels 0;	Gaps 0;
624	TGATGTGGTAGCCATTTCACATCTCTTTGAAAGCTAACAAAGGCACTAAAAATTCGATCGTG	683		
237	TACAAGTAGTTTCACATTTCCATTTTTCGAAGCAATATGCAATTAGATTTGATCGTG	178		
684	AAAAAGCCTATAGCAAAAGCCTATAGATTTCCCTCTGGCAACACGCTACGCATTTGGGGCAG	743		
177	AGAAAGCATATGATAAACAATTTGGATATTTCTGCAGGAGCTGCAGTCAGATTTGAACCTG	118		
744	GACAAACCCGCAAGTGCAGTTGATTCCTCTTGTTGGCAGCTAAAAAAGTCATTTGGCATGA	803		
117	GAGTCAAAAAAAGTACAACTTGTGCAATATTTCTGCACGCTAAAAATTTATGATTC	58		
804	ACGGGCTTGTGAATAAACATCGCGGATGAA	832		
57	GTGGTTTAGTCGATGGCGATATTGACGAA	29		

```

RESULT 15
US-11-266-748A-58311/c
/ Sequence 58311, Application US/11266748A
/ Publication No.: US20060134663A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ TITLE OF INVENTION: Methods of Using the Same
/ FILE REFERENCE: 5815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 58311
/ LENGTH: 1933
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-58311

```

Query Match 1.7%; Score 49.2; DB 8; Length 1933;

```

Best Local Similarity 57.9%; Pred. No. 0.14;
Matches 99; Conservative 69; Indels 2; Gaps 1;

Qy 55 ATTCRCACTGTGTAATRCCTATTATTAATTTTAAATAATTTACTATTATCATATATAATA 114
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1137 ATTATTAAATTATTAAATTGTAATTTATTAAATCTTTAAAGATTATTAAATTATAATTTAAAT 1078

Qy 115 ATATTATTACTTATATTTAAAAAGCTTAATAAAAAAGTAACGAAATTTAGGACATATAATCCCAT 174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1077 ATATTATTAAATCTTAAAGTTATTAAATTATAA - TATTAAATTTAAATTATTAAATCTTT 1020

Qy 175 TGCCTTTAAATTTTAAACAAGAGTAATAATAGTGAACCTCACACCCAAAGA 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1019 AAAGATTATGAATTTAAATTTAAATTTAAATTTAAATTTCTTTAAAGATTAAAGA 969

Search completed: August 10, 2006, 01:50:58
Job time : 447 secs

```

Search completed: August 10, 2006, 01:50:58
Job time : 447 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2006, 01:16:43 ; Search time 12167 Seconds
(without alignments)
13250.221 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2880.6
Sequence: 1 rgragattttccaractt.....aaaaaagtagaacacagg 2883

Scoring table:

IDENTITY_NUC
Gapop 3.0 , Gapext 3.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C					
1	364.8	12.7	956	14	DU750758 ASNF1644.
2	346.2	12.0	936	14	DU760951 ASNG3437.
3	335.2	11.6	942	14	DU762934 ASNG641.g
4	333.2	11.6	958	10	DR637438 EST102806
5	330.2	11.5	1509	6	BR816461 Arabidops
6	330	11.5	946	10	DR637093 EST102771
7	320	11.1	809	10	DV133265 CV03067A1
8	312.2	10.8	925	10	DR633422 EST102404
9	309	10.7	872	8	CO122599 GR_EB04F
10	306.8	10.7	889	10	DV759139 PchrSEQ10
11	302.2	10.5	748	9	CK676965 ydd29H04.
12	300.2	10.4	889	10	DR637386 EST102801
13	296.4	10.3	796	10	DR633939 EST102456
14	293	10.2	849	10	DR633599 EST102622
15	284	9.9	784	11	AZ935182 BJ_Ba000
16	279.6	9.7	766	8	CV195867 CGF100344
17	276.4	9.6	845	10	DV135585 CV03111B2
18	272.4	9.5	784	10	DV758961 PchrSEQ87
19	264.6	9.2	713	5	CK757406 atr02-4ms

C	20	264.4	9.2	986	5	CF826586	EST703968
C	21	262	9.1	946	9	DN809412	73954613
C	22	261.2	9.1	684	9	DN907928	030109ABL
C	23	261	9.1	921	5	CF714848	CCAF343TO
C	24	260.8	9.1	945	9	DN808832	76945303
C	25	259.8	9.0	863	5	CD375065	PTMM00226
C	26	253.2	8.8	905	4	CA264652	SCQGLB204
C	27	252.4	8.8	796	5	CF708414	CCABR70TR
C	28	250.8	8.7	689	10	DR440006	EST1149_09
C	29	249.4	8.7	757	3	BU027777	QHG7L07.Y
C	30	248.6	8.6	781	10	DR916924	EST110846
C	31	248.4	8.6	739	10	DV152796	CV03090A1
C	32	247.8	8.6	661	8	DN908212	030109ABL
C	33	243.2	8.4	789	4	CA236757	SCQF1505
C	34	240.4	8.3	692	8	CV902222	PD019F1 m
C	35	239.8	8.3	937	9	DN810553	CV03090A1
C	36	239.4	8.3	990	8	CO027797	EST806181
C	37	239	8.3	643	14	AG940016	Drosophil
C	38	237.2	8.2	634	2	BG887260	EST751311
C	39	236.8	8.2	1275	12	BZ572656	msb2_2738
C	40	235.8	8.2	660	11	AZ935008	BJ_Ba000
C	41	235.4	8.2	829	5	CF689903	CCADS40TF
C	42	234.8	8.2	636	4	CA851837	D18B02.D1
C	43	234.2	8.1	626	8	CN910029	030124ABL
C	44	231.8	8.0	614	11	AZ935371	BJ_Ba000
C	45	227.4	7.9	747	5	CF690511	CCAAZ91TR

ALIGNMENTS

RESULT 1
DU750758/c
LOCUS DU750758 956 bp DNA linear GSS 27-JAN-2006
DEFINITION ASNF1644.b2 HF130_10-06-02 uncultured marine microorganism
HF130_10-06-02 genomic clone HF0130_019D02, genomic survey
sequence.

ACCESSION DU750758
VERSION DU750758.1 GI:85760594

KEYWORDS GSS.
SOURCE uncultured marine microorganism HF130_10-06-02
ORGANISM uncultured marine microorganism HF130_10-06-02
unclassified sequences; environmental samples.

REFERENCE 1 (bases 1 to 956)
AUTHORS DeLong, E.P., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J.,
Frigaard, N.U., Martinez, A., Sullivan, M., Edwards, R., Chisholm, S.W.
and Karl, D.M.

TITLE Comparative genomics reveals ecological trends in stratified
microbial communities in the ocean's interior

JOURNAL Science (2006) In press
COMMENT Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
and Edward DeLong

US DOE Joint Genome Institute
US DOE Joint Genome Institute
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Tel: 617-253-5271
Fax: 617-253-2679

Email: PMRichardson@lbl.gov; delong@mit.edu
North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
DNA library prepared from marine picoplankton in the less than 1.6
um, greater than 0.22 um fraction. Sample Date: 10/6/2002

Coordinates: 22.45 N, 158 W Depth 130 m Temperature: 22.19 C
Salinity: 35.31 psu Oxygen: 204.9 umol/kg
Class: fosmid ends.

Location/Qualifiers
1. .956
/organism="uncultured marine microorganism HF130_10-06-02"
/mol_type="genomic DNA"
/db_xref="taxon:361146"
/clones="HF0130_019D02"
/cell_type="marine picoplankton, less than 1.8 um, greater
than 0.22 um fraction"

FEATURES
source

/clone_lib="HF130_10-06-02"
/note="vector: pC1FOS; North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid DNA library prepared from marine picoplankton in the less than 1.6 um, greater than 0.22 um fraction. Picoplankton collected at 130 m depth on 10/6/2002, Coordinates: 22.45 N, 158 W. Sample Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 130 m Temperature: 22.19 C Salinity: 35.31 psu Oxygen: 204.9 umol/kg"

ORIGIN

Query Match 12.7%; Score 364.8; DB 14; Length 956;
Best Local Similarity 61.8%; Pred. No. 2.6e-75;
Matches 588; Conservative 0; Mismatches 362; Indels 1; Gaps 1;

1250 AGCACTAGCAGGGAAGGTATGATTATACCGCTGGGGGAATCGATTACACACCCACTT 1309
1310 CTTTCTCCACAAATTCCTACCGCTCTAGCCAATGGCGTTTACCAACCATGTTGGAGG 1369
896 AATCTGTCTCTCAACAAATAGAACTGCTTTAGCTAGTGGGGTTACCAATGCTCGAGG 837
1370 CGGCACAGTCTCTAGATGGCAGCAATGCGACTACTATCACTCCGGCAATGGAATT 1429
836 AGGCACAGGCGCCAGCAACAGGTACTAATGCAACGACATGTACACCTGGAGCATTTACAT 777
1430 GCACCGCATGTTGGCGCAGCAGAAAGAGTATTCATGAATGTGGGCTTTTGGGCAAGG 1489
776 CTCAAGATGCTGCAATCTCGAGAGGATTTCTTGTAAATTTGGGATTCCTTTGGAAGG 717
1490 CAATAGCTCTAGCAAAAACAACTTTAGAACAAAGTAGAGCGGCGCGATTGGTTTAA 1549
716 CAATGCAACTAACAAAGCAGCATTTAGAAAGCAAGTAAGCAGGTCTTGTGGTTAA 657
1550 ATTGCATGAAGATGGGGCAACACCAAGTGCATCGATCATCTGTGAGCGTGGCAGA 1609
656 ACTCATGAAGACTGGGGCAACACCGCGCTGTATTGATTCCTGCTTAACTGTTGCGAGA 597
1610 TGAATAGATGTCGAAGTTTGTATCCACCGATACAGTCAATGAGCAGGCTTATGTA 1669
596 TCAACTAGAGCTCAAGTTTGTATTCATACAGATACCCCTAAATGAAGCTGTTTGTGA 537
1670 TGACACCTAAATGCAATGAACGGCGCGCATCCATCCATCCATACCACTGAGGAGCGG 1729
536 AGATACAAATTAAGCANTAGAGGAGGAGCAATTCATCTCCATACAGAGGGCTGG 477
1730 TGGAGGACACTCACCTGATGTTATCAACATGGCAGGCGAGCTCAATATTTACCCCTC 1789
476 AGGTGGTCAAGCTCCGACATTTATAAAATTTGTGGAGAAATCAAAATGTGATTTCCAGCAG 417
1790 CACCAACCCCACTATTCCTTATACCATTAATACGGTTGCGAGAACTTAGACATGCTCAT 1849
416 TACAAATCCAACTAGGCGCTTCACTCTAAATPACTCTTGAAGAGCATTTAGACATGTTGAT 357
1850 GACATGCACACCACTAGACAAACGCATCCGGAGGATTTACAATTTCTCAAGCCGAT 1909
356 GGTGTGTCATCAATTTAGATCCCAAAATTCAGAGATGTTGCATTTGCTGAGTCAAGAT 297
1910 CCGCCCGGCTCTATCGCGGTGAAGATGTTGCTCCATGATATGGGTGTGATCGCGATGAC 1969
296 ACCTCGTGAACTATTGCTGTGAGGACATCTCCAGGACTTAGAGCGCTTTCTATTAT 237
1970 AAGTCGGATTCGCAAGCAATGGGCGGTGACGGCAAGTGTATTCCTCGAATTTGGCAGAC 2029
236 TGCTAGTGACTCCAGGCTATGGGTAGAGTTGGGGAGGTGATAGCCGGAATTTTCAAC 177
2030 TCGGGAATGAATAAAGATTTTGTAACTTCTTGAAGATGCAAGATGAACGATAA 2089
176 TGCTCATAAATGAAGTTTCAAGAGGAGCGCTTACCTGAGGATATCAAGGAATGATAA 117
2090 TTTCCGATTAAGCGCTTACATCTCTCAATATCACTATCAACCCCGCTTTGACCCAGCGGT 2149

Db 116 TCATGCTCTGAAAAGATATATCTAAAGTCACTATTATTAATCTCGCATAGCTCATGGAAT 57
Qy 2150 GAGCAGATATATCGCTCTGTGGAAGAGGCGCAAGATCGCGACTTGGTGT 2200
Db 56 CAGTGTCTATGTTGTTGCGTAGAG-TTGAAAACCTAGCAGACTTAGTGTAT 7
RESULT 2
DU760951
LOCUS
DEFINITION ASNG3437.92 HF200_10-06-02 uncultured marine microorganism
HP200_10-06-02 genomic clone HF0200_094E10, genomic survey
sequence.
ACCESSION DU760951
VERSION DU760951.1 GI:85770787
KEYWORDS GSS.
SOURCE uncultured marine microorganism HF200_10-06-02
ORGANISM uncultured marine microorganism HF200_10-06-02
unclassified sequences; environmental samples.
REFERENCE 1 (bases 1 to 936)
AUTHORS DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J.,
Frigaard, N.U., Martinez, A., Sullivan, M., Edwards, R., Chisholm, S.W.
and Karl, D.M.
TITLE Comparative genomics reveals ecological trends in stratified
microbial communities in the ocean's interior
JOURNAL Science (2006) In press
COMMENT Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
and Edward DeLong
US DOE Joint Genome Institute
US DOE Joint Genome Institute
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Tel: 617-253-5271
Fax: 617-253-2679
Email: P.M.Richardson@lbl.gov; delong@mit.edu
North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
DNA library prepared from marine picoplankton in the less than 1.6
um, greater than 0.22 um fraction. Sample Date: 10/6/2002
Coordinates: 22.45 N, 158 W Depth 200 m Temperature: 18.53 C
Salinity: 35.04 psu Oxygen: 198.8 umol/kg
Class: fosmid ends.
Location/Qualifiers
1..936
/organism="uncultured marine microorganism HF200_10-06-02"
/mol_type="genomic DNA"
/db_xref="taxon:361147"
/clone="HF0200_094E10"
/cell_type="marine picoplankton, less than 1.8 um, greater
than 0.22 um fraction"
/clone_lib="HF200_10-06-02"
/note="Vector: pC1FOS; North Pacific Subtropical Gyre
(Hawaii) picoplankton genomic fosmid DNA library prepared
from marine picoplankton in the less than 1.6 um, greater
than 0.22 um fraction. Picoplankton collected at 200 m
depth on 10/6/2002, Coordinates: 22.45 N, 158 W. Sample
Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 200 m
Temperature: 18.53 C Salinity: 35.04 psu Oxygen: 198.8
umol/kg"

ORIGIN

Query Match 12.0%; Score 346.2; DB 14; Length 936;
Best Local Similarity 60.8%; Pred. No. 6.9e-71;
Matches 564; Conservative 0; Mismatches 363; Indels 0; Gaps 0;

Qy 1665 GTAGATGACACCTAAATGCAATGAACGGCGCGCATCATGCTCCCTACCACTTAGGGA 1724
Db 2 GTGAGAAATACCGTCGCGCATGAAGCAGCGCTCATCCACGCTTCCATCTAGAGGG 61
Qy 1725 GCGGTGGAGGACACTCACTGATGTTATCCATCGCAGGCGAGCTCATATTTCTACCC 1784
Db 62 GCGGTGGCGGCAACGCGCGGACATCAAGATCTGCGGGAAGCCATGTGTCGCC 121
Qy 1785 TCCTCCACCACCCCACTATTCCTATACCATTAATACGTTGCGAAGACACTTAGACATG 1844

```
Db 122 TCATCGACCAACCGCGCGCCCTACACGCTGAACAGCTTGAAGCATCTCGACATG 181
Qy 1845 CTATGACATGCCACCTAGACAAACGATCGCGAGGATTTACAAATTTCTCAAGC 1904
Db 182 CTATGCTCTGCCATCACTCGACAAGTCAATACCGGAAGATGTGGCTTTGCGGAAGC 241
Qy 1905 CGPATCGCCCGCGCTCTATCGCGGCTGAGATGCTCCATGATATGGGTGATCGCG 1964
Db 242 CGATCAGGCGGGAACATCGCGGAGAGACATCTCTTATGACATGGGGGCTTTTCG 301
Qy 1965 ATGACAACTCGGATTCGCAAGCAATGGGCGTGCAGCGAAGTATCTCTCGAACTTGG 2024
Db 302 ATCATCGGCTCGACACCGAGGCATGGCGGTCTCGCGAGGTTCATCATCGGACCTGG 361
Qy 2025 CAGACTCGGATAGAAATAAAGAAATTTGGTAAGCTTCTGGAAGTGGCAAGATAAC 2084
Db 362 CAGACAGCGCAAGATGAAGTCCAGCGCGAAGGCTGCTGAGGAACCGGCGCAAC 421
Qy 2085 GATAATTTCCGATTAAGCGCTACATCTCCAAATACATCACTATCAACCCGCTTTGACCCAC 2144
Db 422 GACAATTTGCGGGTCAAGCGCTATGTGCGAATATACCACTCAACCCGCGCATCGCCAT 481
Qy 2145 GCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTGTGTGG 2204
Db 482 GGATATCAGCGCATATCGGGTCGATCGCGTGGGAAGCGGCGGACCTTGTGCTCTGG 541
Qy 2205 AATCTCTGCTTTTGGCGTAAACCCAAATCGTGATCAAGGCGGTATGGTGTCTTC 2264
Db 542 AATCGGCTCTTTCGGCGTGAAGCGGAGATGTCATCATCGGCGGCGAGCATTCGCTGC 601
Qy 2265 TCTGAATGGCGATTTAAGCGCTGTGTGCCACTCCCAACCGGTTTATACCGGAA 2324
Db 602 GCGCAGATGGCGATCTTAAGCGCTCAATCCGACACCGGAGCTGTCTACACCGCGCG 661
Qy 2325 ATGTTTGGGATCAAGCGAGGGAATTTGACACAGCATCTTTTGTTCAAAGTC 2384
Db 662 ATGTTGCGCGGCTTTCGGCAGGTGGTGAGAAATTCGGCTGTCACTTCGTCAGCGCAAG 721
Qy 2385 GCCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGGTCAAA 2444
Db 722 GCGCAGATGCGCGATCGCGGACACGCTCGGCTTGCAGGGGACCGTCCGCGTGGCG 781
Qy 2445 AACTGCCGTAACATCAACGAAGAAGCTTCAAGTTCAACGACAAAGCGCAAAATCAAC 2504
Db 782 AAGACGCTACGATCTCCAGGCGGACATGTTGCATAACGATTACTGCCCGGACATCGAG 841
Qy 2505 GTCGATCCGAAACCTTCGAGGCTTTGTAGATGGCAAACTCTGCACTCTAAACCCACC 2564
Db 842 GTCGACCCGAAACCTATGAGGTTTCGCGCTGATGGCGAGCTTCTGACCTCGAGCGCGG 901
Qy 2565 TCGCAAGTGCCTCTAGCCCGACGCTAC 2591
Db 902 GAAGAACTGCGGATGGCGAGCGCTAC 928
```

RESULT 3

```
LOCUS DU762934
DEFINITION 942 bp DNA linear GSS 27-JAN-2006
ASN3641.g2 HF200_10-06-02 uncultured marine microorganism
HF200_10-06-02 genomic clone HF0200_064A09, genomic survey
sequence.
ACCESSION DU762934
VERSION DU762934.1 GI:85772770
KEYWORDS GSS.
SOURCE uncultured marine microorganism HF200_10-06-02
ORGANISM uncultured marine microorganism HF200_10-06-02
unclassified sequences; environmental samples.
REFERENCE 1 (bases 1 to 942)
AUTHORS DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
and Karl,D.M.
TITLE Comparative genomics reveals ecological trends in stratified
```

microbial communities in the ocean's interior

JOURNAL

Science (2006) In press
Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
Kerrie Barry, Tjiana Glavinadelrio, David Bruce, Paul Richardson
and Edward DeLong

COMMENT

US DOE Joint Genome Institute
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Tel: 617-253-5271
Fax: 617-253-2679

Email: PWRichardson@lbl.gov; delong@mit.edu

North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
DNA library prepared from marine picoplankton in the less than 1.6
um, greater than 0.22 um fraction. Sample Date: 10/6/2002
Coordinates: 22.45 N, 158 W Depth 200 m Temperature: 18.53 C
Salinity: 35.04 psu Oxygen: 198.8 umol/kg
Class: fosmid ends.

FEATURES

Location/Qualifiers

1..942

/organism="uncultured marine microorganism HF200_10-06-02"
/mol_type="genomic DNA"
/db_xref="taxon:361147"
/clone="HF0200_064A09"
/cell_type="marine picoplankton, less than 1.8 um, greater
than 0.22 um fraction"
/clone_lib="HF200_10-06-02"
/note="Vector: pCC1FOS; North Pacific Subtropical Gyre
(Hawaii) picoplankton genomic fosmid DNA library prepared
from marine picoplankton in the less than 1.6 um, greater
than 0.22 um fraction. picoplankton collected at 200 m
depth on 10/6/2002, Coordinates: 22.45 N, 158 W. Sample
Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 200 m
Temperature: 18.53 C Salinity: 35.04 psu Oxygen: 198.8
umol/kg"

ORIGIN

Query Match 11.6%; Score 335.2; DB 14; Length 942;
Best Local Similarity 62.6%; Pred. No. 2.8e-68;
Matches 538; Conservative 0; Mismatches 318; Indels 3; Gaps 1;
Qy 1308 TTCCTTTCTCCACAAATTCCTACCGCTCTAGCCATGGCGTTACAACTATGTTGA 1367
Db 1 TTTATCTCGCCCAATCTTATTTCTGAAGCTTCTATTCGGCATCACCACTCATCTGC 60
Qy 1368 GCGCGCACAGTCTCTAGATGGCAAGATGCGACTATCTACCTCCGGGCAATGGAAC 1427
Db 61 GGTGGCAGGGGCGGCGGACTGGCACCAACGCGACCACTGCACTCCGCGTGGGAC 120
Qy 1428 TTGCACCGCATGTTGCGCGCAGCAGAGATTTCTATGAATGTGGGCTTTTGGGCAAA 1487
Db 121 CTCGCAAAATGTACGAGGCGGTGGAGCGTTCCCGCTCAACTTCGGCTTTTGGGCAAG 180
Qy 1488 GGCATAGCTCTAGCAAAACAACTTGTAGAACAGTAGAGCGGCGGATGGTTT 1547
Db 181 GGCACAGCTCGCTTCCAAATGCCCTTGGGGAACAGGTGGAAGCGGAGCGATTGGCCTG 240
Qy 1548 AAATTGCATGAAGACTGGGGCACAAACCAAGTCGATCGATCTGTGACGCTGGCA 1607
Db 241 AAATTGCAGAGACTGGGGCACAGCTCCGGGGCGATGTACTGTCTCGCGTGGCG 300
Qy 1608 GATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGCAGGTTATGTA 1667
Db 301 GAGGAGTACGAGTGCAGGTGCGCATCCACACCGACCGCTCAACGAATTCGGCTTTGTG 360
Qy 1668 GATGACACCTTAATGCAATGAAGCGGCGCGCATTCATGCTTACACATTCAGGAGCG 1727
Db 361 GAGGACACGCGGGCGGCTTCAAGGAGCGCACCATCCACACCTACCACTGAGGCGCA 420
Qy 1728 GGTGGAGGACCTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCTCC 1787
Db 421 GCGGCGGACACGCGCCCGGACATCATCAAGCTCTCGGCGAGGCGCCAGCTGCTGCTCC 480
Qy 1788 TCCACCAACCCCACTATTTCCTATACCAATTAATACGTTTGCAGAACACTTAGACATGCTC 1847

```
Db 481 TCACCAATCCGACGATGCCGTTACCGCAACACGATGACGAGCATCTGGACATGCTG 540
QY 1848 ATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAAATTTCTTCAAGCCGT 1907
Db 541 ATGGTCTGCCACCACTCTCGCCCGCATTCGCGAGGACGTCGCTTTGCGATCGGC 600
QY 1908 ATCCGCCCGCGCTTATPCGCGGCTGAAGATGTCTCATATGATATGGGTGATCGCGATG 1967
Db 601 ATCCGCCCGGAGACCAATTCGCGCGGAGACATCTGCACGACCTCGCGCGGCTTCAGCAAT 660
QY 1968 ACAAGCTCGATTCGCAAGCAATGGGCGCTGCAGGCAAGTATCTTCGAACTTGGCAG 2027
Db 661 CTGCGCTCGAATTCGCAAGCGATGGCGCGCATCGCGCAAGTGAATTCGCGCACCTGGCAG 720
QY 2028 ACTGCGGATAAGAAATAAAAGAAATTTGTAAGCTTCCTGAAGATGCAAAAGATAACGAT 2087
Db 721 ACCGCCGCAAGATGAAGCCGACGCGGCTCGCTTCGCGGCGAGCGG---GGAGCGAC 777
QY 2088 AATTTCCGCAATTAAGCGCTACATCTCAAAATACACTATCAACCCCGCTTTTGACCCACGGC 2147
Db 778 AACTTCGCATCCGTCGCTACGTCGCGCAATACACCATCAACCCCGCATCGCCCGCAGCGC 837
QY 2148 GTGAGCGAGTATATCGGCT 2166
Db 838 ATAAACAAGGTCGTTGGCT 856

RESULT 4
DR637438
LOCUS
DEFINITION
DR637438 958 bp mRNA linear EST 11-JUL-2005
EST1028063 FvM Gibberella moniliformis cDNA clone FVMAW64, mRNA
sequence.
ACCESSION
DR637438
VERSION
DR637438.1 GI:70712272
KEYWORDS
Gibberella moniliformis
ORGANISM
Gibberella moniliformis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 958)
Brown,D.W., Cheung,F., Proctor,R.H., Butcho,A.E., Zheng,L, Lee,Y.,
Utterback,T., Smith,S., Feldblyum,T., Glenn,A.E., Plattner,R.D.,
Kendra,D.F., Town,C.D. and Whitelaw,C.A.
Comparative analysis of 87,000 expressed sequence tags from the
fumonisin-producing fungus Fusarium verticillioides
Fungal Genet. Biol. 42 (10), 848-861 (2005)
16099185
Contact: Brown, D.W.
USDA/ARS/NCAUR
USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncaur.usda.gov
TIGR sequence name: FVMAW64TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
Location/Qualifiers
1..958
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAW64"
/tissue_type="mycelia"
/clone_lib="FvM"
/notes="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to TRIzol Reagent (Invitrogen, Carlsbad CA) at
```

```
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."
```

ORIGIN

Query Match 11.6%; Score 333.2; DB 10; Length 958;
Best Local Similarity 60.6%; Pred. No. 8.5e-68;
Matches 572; Conservative 0; Mismatches 368; Indels 4; Gaps 2;

QY 1209 GATCGCGTAAGCCCTCATATGCTCGTGGTGTGGGCACAGAAAGCACTAGCAGGGGAAGGT 1268
Db 2 GACCGCGTAAACAGAGGCGATGCTCGGAAGCTGACGGATGTCGTACAGGTGAAGGA 61
QY 1269 ATGATTAATTACCGCTGGGGGAATCGAATTCACACCCCACTTCTCTTCTCCACAACAATTC 1328
Db 62 AAGATTGTGACCGCGGGCGCTATCGACACGCATATTCACTTTATCTGCCCTCAGCAAGTA 121
QY 1329 CTTACCGCTCTAGCCAATGGCGTTACACCATGTTTGGAGGCGGCACAGTCTCTAGAT 1388
Db 122 CCCGAAGCTCTTGCATCTGCTGTAAACCAACATGCTTGGCGGTGGTACCGGCCCAAGTGCT 181
QY 1389 GGCACGAATCGACTACTATCACTCCGGGCAATGGAACCTTCACCGCATGTTTGGCGCA 1448
Db 182 GGAACGAACGCAACTTGTACGCTGTGCTCATTAATGCTCAATGTTTGCAGGCG 241
QY 1449 GCAGAAGAGTATTTCTATGAATGTGGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAA 1508
Db 242 TCGCATCAGCTTCTTATCAATATTGTTACTTGGCAAGGTAATGATAGTTCTCTCTGAG 301
QY 1509 CAATCTGTAGAACAAAGTAGAGCGGGCGGATGGTTTTTAAATTTGATGAAGACTGGGGC 1568
Db 302 GGTCTGCCGATCAGGTCAATGCTGTGCTTGGCTCAAGCTTCAAGGACTCGGGGT 361
QY 1569 ACAACACCAAGTGGCATCGATCAGTCTTGGCGTGGCAGATGAATACCATGTGCAAGTT 1628
Db 362 TGCACTCTCTGTCTATTGACGCTTGTCTCAGTGTCTGTGATGAATTCGATATCAATGT 421
QY 1629 TGTATCCACACCGATACAGTCAATGAGGCGAGTATGTAGATGACACCCCTAAATGCAATG 1688
Db 422 CTTATTACACTGACACGCTTAAACAGTGTGCTTGTGCGAATCTACGATCGCTGTCTTC 481
QY 1689 AACGGCGCGGCATCCATCGCTTACACATTTAGGAGCGGGTGGAGGACACTCACCTGAT 1748
Db 482 AAGAACCCGCAAAATTCATCTTATCACACAGAGGGTGCAGGAGGTGCCCATGCTCCGGAT 541
QY 1749 GTTATCACATGGCAGCGAGCTCAATATTCTACCTCTCCACACCCGCCATATTTCCC 1808
Db 542 ATCATCTCCGTGGTAGAGCATCAAAATGTTCTGCCATCATCGACCAACCCCTCAAGACCA 601
QY 1809 TATACCATTAATACGGTTGCAAGACACTTAGACATGCTCATGACATGCCACCATCTAGAC 1868
Db 602 TTCACACGCAATCTCTCGATGAGCATCTCGATATGCTTATGGTCTGCCATCACTTGTCC 661
QY 1869 AACCGCATCCGCGAGGATTTACAAATTTCTCAAAGCGGTATCCGCCCGGCTTATCGCG 1928
Db 662 AAGAATAATCCAGAGGATGTAGCCCTTCGCGGAGAGCGGTATTTCGTGCTGAACCAATTGCT 721
QY 1929 CTTGAAGATGTCTTCATGATATGGTGTGATGCGGATGACAGCTCGGATTCGCAAGCA 1988
Db 722 GCTGAGGATGTATTAACGACAAAGGCGCTATCAGCATGTAGTCTCTCACTCGCAGGCT 781
QY 1989 ATGGGGCGTGCAGCGGAAGTGAATTCCTCGAATCTGCGACACTGCGGATGAAGATAAAAAA 2048
Db 782 ATGGCCCGTTCGGAGAGGTCGTTTTAAGAACATGGAATATCTGCGATTAATAATAGGTG 841
QY 2049 GAATTTGGTAAGCTTCTCTGAAGAT---GGCAAGATAACGATAATTTCCGCATTAAGCGC 2105
Db 842 CAGAGGGGTGTTTCGCGGAGGATGAAGGSCACAGGGGCTGATAATGCGCG-TGTAAACGC 900
QY 2106 TACATCTCAAAATACACTATCAACCCCGCTTTGACCCACGGGT 2149
Db 901 TATGTGACGAAGTATACTATTAAACCCAGCTATTGTCTCAGGCTT 944

RESULT 5
 CNSOABR8 1509 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSLTPGH42D08 of Hormone Treated Callus of strain col-0 of
 Arabidopsis thaliana (thale cress).
 ACCESSION BX816461.1 GI:42471655
 VERSION HTC; GSLT cDNA.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1509)
 AUTHORS Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1509)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV-INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.
 FEATURES
 Location/Qualifiers
 1..1509
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="GSLTPGH42D08"
 /tissue_type="Hormone Treated Callus"
 /ecotype="Col-0"
 /plasmid="pCMVSPORT_6"
 1..1509
 /gene="At1g67550"
 gene
 1..1509
 /gene="At1g67550"
 ORIGIN
 Query Match 11.5%; Score 330.2; DB 6; Length 1509;
 Best Local Similarity 59.1%; Pred. No. 5e-67;
 Matches 602; Conservative 0; Mismatches 413; Indels 4; Gaps 4;
 QY 1587 GATCACTGCTTGAGCGGTGCAGATGATGCAAGTTTGTATCCACCGCATACA 1646
 DB 1 GACAAATGTTGGCAGTTGGCAGATATGATATCCAGGTGAACATTCATCTGACACC 60
 QY 1647 GTCAATGAGCAGGTTATGATGATGACACCTTAATGAATGAACGGCGCGGCATCAT 1706
 DB 61 TTAATGAATCCGGTTTGTGTAGCACAATCAATCAATTCGTGGGAGAACAAATACAT 120
 QY 1707 GCCTACCAATTCAGGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGC 1766
 DB 121 ACATATCACAGTAAGGTGCTGGTGGGACATGCACCATATCATATAGAGTTGTGGA 180
 QY 1767 GAGCTCAATATTCCTCCTCCACCCACCTATTCCTTATACCATTAATACCGTT 1826

Db 181 GTGAAATGTACTCCCGTCATCAACCAACCCCAACGCCATATACGAAATACTGTA 240
 QY 1827 CGAGAACACTTAGACATCTCATGATGACATGCGCACCACTAGACAAACGCATCCCGGAGGAT 1886
 Db 241 GATGAACATCTTGACATGCTGATGTTGGCCATCACCCTTGACAAGAACATCCCGAGAAGAT 300
 QY 1887 TTACAAATTTCTCAAGCGGTATCCGCCCGGCTCTATCGCGGCTGA-AGATGTGCTCCA 1945
 Db 301 GTAGCTTTTGTGAATCAAGATAAGAGCTGAAACAATAGCTGCAGAGAGATATATTGCA 360
 QY 1946 TGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGCTGCAGCGCA 2005
 Db 361 TGATATGGAGCAATCAGCAATTTCTCTCCGACTCACAAGCAATGGGTGCGATTGGAGA 420
 QY 2006 AGTGATTCCTCGAACTTCGCGAGCTGCGGATGAAGATAAAAAAGAAATTTGGTAAGCTTC 2065
 Db 421 GGTAATAAGCAGAACTTGGCAGACAGCTGACAAGATGAAGGCTCAGAGGGGAGCCATTA 480
 QY 2066 TGAAGATGGCAAGATAAGATAATTTCCGCATTAAGC-GCTACATCTCAAATACACTA 2124
 Db 481 TCCTAACATGGCTGATGATGATACTCAAGGATCAAGCAGATACATTGCAAGATGACACTA 540
 QY 2125 TCACCCCGCTTTGACCCACCGCGTGAGGAGTATATCGCTCTGTGGAAGAGGGCAAGA 2184
 Db 541 TAAACCCAGCTATAGCCAAATGGGTTTGAGACTTAATTCGCTCGGTTGAGGTGAAGAAGC 600
 QY 2185 TCGCGCACTTGGTGTGTGGAATCC-TGCCTTTTTTGGCGTAAACCCAAAATTCGTGATC 2243
 Db 601 TGGCTGATCTTGTATATATGCGCAGCGCGCTTCTTCGGAGCAAAACCCAGAAATGATAATC 660
 QY 2244 AAGCGGTATGGTGTCTTCTGAAATGGCGATTTTAAACGGTCTTGTGCGCACTCC 2303
 Db 661 AAGGAGCAATATAGCATGGCTAATATGGCGATGCAAAATGCGAGCATACCCACTCT 720
 QY 2304 CAACGGTTTATTACCGCGAAATGTTGGGCATCAGCGCAAGCGGCAAAATTTGACACCAGC 2363
 Db 721 GAGCGGTGATATCAAGACTTATGTTGGAGCAATTTGGAAGCGCGGAAGTGAAGAACTCT 780
 QY 2364 ATCACTTTGTTTCCAAAGTCCTTATGAAATGGCGTGAAGAAAGTGGGCTTAGAG 2423
 Db 781 GTTCATTTGTCAGGAGGCTGCTTGAAGAGGGGGTAAAGAACTATATGATGATCAAG 840
 QY 2424 CGCCAAGTTCTACCGGTCAAAAACCTGCGTAAACATCACCAGAAAGACTTCAAGTTCAAC 2483
 Db 841 AAGAGGGTTGTAGCGGTGTCCAACTGAGGAGCTCACAATACTCCACATGAAGCTGAAT 900
 QY 2484 GACAAACGGCAAAATCACCCTGATCCGAAACCTTCGAGGTCTTTGTAGATGCAAA 2543
 Db 901 GATCGGCTTCCAGAGATCACCCTGATCCAGAGATACGTTGTTCACAGCAAAATGCGGAG 960
 QY 2544 CT-CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCACGCGTACACTTTCTTCT 2601
 Db 961 GTCTTAGCATGTGGCGCAGCGGATTCAGTCCCTCTCTGCGCGCGGAACATTTCTCTCT 1019

RESULT 6
 DR637093
 LOCUS EST1027718 FvM Gibberella moniliformis cDNA clone FvM95, mRNA
 DEFINITION sequence.
 ACCESSION DR637093
 VERSION DR637093.1 GI:70711927
 KEYWORDS EST.
 SOURCE Gibberella moniliformis
 ORGANISM Gibberella moniliformis
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 REFERENCE 1 (bases 1 to 946)
 AUTHORS Brown, D.W., Cheung, F., Proctor, R.H., Burchko, A.E., Zheng, L., Lee, Y.,
 Unterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
 Kendra, D.F., Town, C.D. and Whitelaw, C.A.
 TITLE Comparative analysis of 87,000 expressed sequence tags from the
 fumonisin-producing fungus *Fusarium verticillioides*

JOURNAL Fungal Genet. Biol. 42 (10), 848-861 (2005)

PUBMED 16099185

COMMENT Contact: Brown, D.W.

USDA/ARS/NCAUR

USDA

1815 N. University St, Peoria, IL 61604, USA

Tel: 309 681 6230

Fax: 309 681 6689

Email: brown@ncaur.usda.gov

TIGR sequence name: FVMT95TH

Seq primer: AAT TAA CCC TCA AAG GG.

Location/Qualifiers

1..946

FEATURES

source

/organism="Gibberella moniliformis"

/mol_type="mRNA"

/strain="m3125"

/db_xref="taxon:117187"

/clones="FVMT95"

/tissue_type="mycelia"

/clone_lib="FVM"

/note="Vector: pBlueScript II SK(+). Site 1: EcoRI; Site 2: XhoI; anamorph: Fusarium verticillioides. Library FVM was prepared from pooled RNA obtained from a 48-hour and a 72 hour, liquid GYM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were frozen in liquid nitrogen, ground to a powder, and then added to Trizol Reagent (Invitrogen, Carlsbad CA) at approximately 1 g mycelia per 10 ml Trizol. The cDNA was directionally ligated into the pBlueScript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

Query Match 11.5%; Score 330; DB 10; Length 946;
Best Local Similarity 60.4%; Pred. No. 4.9e-67;
Matches 570; Conservative 0; Mismatches 370; Indels 4; Gaps 2;

QY 1209 GATGGCTAAGCCCTCATATGGTGTGGTGTGGGACAGAGCACTAGCAGGGGAAGGT 1268
DB 2 GACGCGGTAAACAGAGGCGATGCTGCGAAGCTGACGGATGCTGTAGCAGGTGAAGGA 61
QY 1269 ATGATTATTACCGTGGGGATGATTCACACACCACTTCCCTTCCACACAAATTC 1328
DB 62 AAGATTGTGACCGGGGGCGCTATGACACGATATTCATTTATCTGCGCTCAGCAAGTA 121
QY 1329 CCTACCGCTTAGCCAAATGGCGTTACACCACTGTTTGGAGCGGCACAGGTCCTGTAGAT 1388
DB 122 CCGAAGCTTGTGATCTGGTGTACCAACATGCTTGGCGTGTGTACCGGCCAAGTCT 181
QY 1389 GGCACGAATGCGACTACTATCACTCCGGGCAAAATGGAACTTGCACCGCATGTTCGCGCA 1448
DB 182 GGAACGAACGCACTACTTGTACGCTGCTCAATTACATGGTCAAAATGTTGCAGCG 241
QY 1449 GCAGAGATATTCTGAAATGTGGCTTTTGGCAAGGCAATAGCTTAGCAAAAAA 1508
DB 242 TCGCATCAGCTTCTTATCAATTTGGTATTACTGSCAAAGGTAATGATGTCTCTGAG 301
QY 1509 CAACCTTGTAGAACAGTAGAGCGGCGGATGTTTAAATTCATGAAGACTGGGCG 1568
DB 302 GGTCTGCGCATCAGTCAATGCTGGTCTGTGGCTCAAGCTTCAATGAGACTGGGGT 361
QY 1569 ACAACACCAAGTGGGATGATCACTGCTTGGCGTGGCAGATGAATACGATGTGCAAGTT 1628
DB 362 TGCATCTCTGCTATTGACGCTGTCTCAGTGTCTGTGATGAATTCGATATTCAATGT 421
QY 1629 TGTATCCACCGATACAGTCAATGAGGCGAGGTATGTAGATGACACCCCTAAATGCAATG 1688
DB 422 CTTATTACACTGACACGCTTAAACGAGTCTGGCTTTTGTGCAATCTACGATCGCTGCTTC 481
QY 1689 AACGGGCGGCTCCATGCTTACACATTCAGGAGCGGTGGAGGACACTCACTCAT 1748
DB 482 AAGAACCGCAAAATTCATCTATACACAGAGGTTGACGAGGTGGCGATGCTCCGGAT 541
QY 1749 GTTATACCATGGGAGGCTCAATATTCTACCTCTCTCCACCACCCCACTATTCCC 1808

DB 542 ATCATCTCCGTGTAGAGCATCAAAATGTTTGGCATCATCGACCAACCTTCAAGACCA 601
QY 1809 TATACCATTAATACGTTTGCAGAACACATTTAGACATGCTCATGACATGCCACCATCTAGAC 1868
DB 602 TTCACACCAATACTCTCGATGAGCATCTCGATATGCTTATGCTCTGCCATCACTTGTTC 661
QY 1869 AAACGCATCCGCGAGGATTTACAAATTTCTCAAACCGGTATCCGCCCGGCTCTATCGCG 1928
DB 662 AAGAATATCCAGAGGATGTAGCTTTCGCGGAGAGCGGTATTCGTGCTGAAACCAATTGT 721
QY 1929 GCTGAAGATGCTCCCATGATATGGTGTGATGCGGATGACAGCTCGGATTCGCAAGCA 1988
DB 722 GCTGAGGATGTATTACACGACAAAGCGCTATCAGCATGATGAGCTCTGACTGCGAGGT 781
QY 1989 ATGGGCGTGCAGGCGAAGTGAATCTCGAACTTGGCAGACTGCGGATGAAGATAAAAAA 2048
DB 782 ATGGGCGTTCGCGAGAGGTGCTTTTAAGAACATGATATCTGCGCATAAAAAATAGTG 841
QY 2049 GAATTTGTAAAGTTCCTGAAGAT--GGCAAAGATAACGATAATTTCCGCATTAAGCGC 2105
DB 842 CAGAGGGGTGTTGTTGCGGAGGATGAGGCGACAGGGGCTGATAATGCGCG-TGTAACGC 900
QY 2106 TACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGCGT 2149
DB 901 TATGTCAGCAAGTATATCTATTACCCCGAGCTATTGCTCAGGGCTT 944

RESULT 7

DVL33265

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Euphorbia esula (leafy spurge)

Euphorbia esula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

Rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;

Euphorbiaceae; Euphorbia.

1 (bases 1 to 809)

REFERENCE

AUTHORS

Anderson, J.V., Horvath, D.P., Thimmapuram, J., Liu, L., Hernandez, A.,

Kim, W.R. and Mikel, M.

Direct submission, Anderson, J.V. 2005

Unpublished (2005)

CONTACT: James V. Anderson

USDA/ARS

1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND

58105, USA

Tel: 701 239 1263

Fax: 701 239 1252

Email: andersjv@fargo.ars.usda.gov

Adaptors and tags in 5'-end sequenced clones:

(Vector) . . . TAAGCTTCATATCG (End Vector) (Start

EcoRI adaptor) AATTCATCTGTGGG (End EcoRI adaptor) (Start

Insert) . . . AAAAAA (End Insert) (Start Tag) TGGCT (End

Tag) (Start NotI site/Vector) GCGCGCCACCCGCGG. . . Base

Calling/Quality Scores: PHRED from Washington University Genome

Center. Vector Trimming: Cross match from Washington University

Genome Center PHRAP suite. Low quality bases (Phred score < 20)

were trimmed from both ends of the sequence by an in-house script.

This sequence is vector free and at least 200 bp in length.

PCR Primers

FORWARD: TAATACGACTCACTATAGG (T7)

BACKWARD: ATTAACCCCTCACTAAAG (T3)

Insert Length: 809 Std Error: 0.00

Plate: CV03067A1 row: A column: 09

Seq primer: TAATACGACTCACTATAGG (T7)

High quality sequence stop: 809.

Location/Qualifiers

1..809

FEATURES

source

```
/organism="Euphorbia esula"
/mol_type="mRNA"
/db_xref="taxon:3993"
/clone="CV03067/Ala09.f1"
/lab_host="DH10B"
/clone_lib="CV03-normalized library"
/note="Vector: pBS II SK(+); Site 1: EcoRI (5' side of insert); Site 2: NotI (3' side of insert); Tissue: Seasonal crown buds from outdoor plants, whole plant tissue (leaf, stem, flowers, seeds, shoots, galls, root and crown buds) from outdoor plants, cold acclimated tissues (leaf, stem, and crown buds), crown and root buds from decapitated greenhouse p ants, and seedlings. The library was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. An identifying tag was added at the 3'during cDNA synthesis: insertAAAAAAAAAAAAAAAAAATGCGT."
```

ORIGIN

Query Match 11.1%; Score 320; DB 10; Length 809;
Best Local Similarity 62.2%; Pred. No. 1.1e-64;
Matches 503; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 1566 GGCAACACCAAGTCGATCGATCCTGTTGAGCGTGCAGATGAATACGATGCGAA 1625
DB 1 GGAATCTCTCTGCTGTATGACAACTGTTTGACGTAGCAGAGAAATATGATCTGCAG 60
QY 1626 GTTTGTATCCACACCGCATACAGTCAATGAGGAGGTTATGTATGATGACACCTCAATGCA 1685
DB 61 GTTAATATCCACACGGACACCTTGAACGAATCCGATTTGTGGAACACACAAATGCTGCA 120
QY 1686 ATGAACGGCGCGCATTCATGCTTACACATTTAGGGAGCGGTGGAGGACACTCACCT 1745
DB 121 TTTAAAGGGAGAACTATACATCTATATCACAGTGAAGGTGCGGAGGTGTCATGCTCCG 180
QY 1746 GATGTTATACCATGGCGGAGCTCAATATTTCTACCTCTCCACACCCCACTATT 1805
DB 181 GACATTATAAAGATGTGTGGTGTAAAAAATGTTATACCGTCATCAACGAATCCTACTCGG 240
QY 1806 CCCTATACCATTAATAGGTGTCAGAACACCTTAGACATGCTCATGACATGCCACCACTTA 1865
DB 241 CCTTTACTTCCATACATCAATGATGACATCTCATATGCTGATGCTGTCGATCATCTG 300
QY 1866 GACAAACGATCCGCGAGGATTTACAAATTTCTCAAGCCGTATCCGCCCGGCTCTATC 1925
DB 301 GATAAGACATTCGGAGATGTAGCTTTCGCTGAGTCACGATAAGGGCGGAACAAATC 360
QY 1926 CGCGCTGAAGATGTGCTCCATGATATGGGTGTATCGGATGACAGCTCGGATTCGAA 1985
DB 361 GCTGCAGAGACATATTGTCATGATATGGGGCTATTAGTATCATTTCTCTGTATTCGAA 420
QY 1986 GCAATGGGGGTGCAGCGGAAGTATTCCTCGAACTTGGCAGCTGCGGATAAGAAATAAA 2045
DB 421 GCTATGGGTGCATTTGGAGAGGTTATATGCAAGACTTGGCAAACTGCTCAAGATGAA 480
QY 2046 AAGAATTTGGTAAGCTTCTCGAAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGC 2105
DB 481 TCACAAGGGGATCAATTTGGTCCGGGCAATCAGATAACGATAATCTCCGATTAACGA 540
QY 2106 TACATCTCCAAATACATATCAACCCCGTTTGAACCAACGCGGCTGAGCGATATATCGGC 2165
DB 541 TACATTTCTAAGTACACTATAATCTCGCATTTAGCAAAATGGGTTTGGCGATTAATTTGGG 600
QY 2166 TCTGTGGAAGGCGCAGATCGCGACTTGGTGTGTGGATCTCTGCTTTTGGGGTA 2225
DB 601 TCAGTTGAGGTAGGAAATTTGCTGATCTTGTATGTGGAAGCGCGGCTTCTTTGGGGCA 660
QY 2226 AAACCCAAATCTGTATCAAGGGGGTATGGTGTCTTCTGAAATGGGGGATCTCTAAC 2285
DB 661 AAACCCGAATGTGTATTAAGGAGGTGTAAATGCTTTGGGCGAGCATGGGTGATCCAAAT 720
QY 2286 GCCTCTGTGCCCATCTCCCCAACCGGTTTATTAACCGCGAAATTTTGGGCATCAACGGCAAG 2345
|||||

721 GCAAGCATCCCAACCCGAAACCGGTGATNTCGAGGCTATGTTTGGAGCATTTTGGCAAG 780

QY 2346 GCGAATTTTGACACACGATCATCTTTTGT 2374
|||

DB 781 GCTGGAAGTGCTAACTCCATTCGTTTGT 809
|||||

RESULT 8
DB633422
LOCUS
DEFINITION
EST1024047 FVM Gibberella moniliformis cDNA clone FVMA534, mRNA
sequence.
ACCESSION
DB633422 GI:70708256
VERSION
DB633422.1
KEYWORDS
Gibberella moniliformis
SOURCE
Gibberella moniliformis
ORGANISM
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
1 (bases 1 to 925)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D., and Whitelaw, C.A.
Comparative analysis of 87,000 expressed sequence tags from the
fumonisin-producing fungus *Fusarium verticillioides*
Fungal Genet. Biol. 42 (10), 848-861 (2005)
16099185
Contact: Brown, D.W.
USDA/ARS/NCAUR
USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncaur.usda.gov
TIGR sequence name: FVMA534TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
Location/Qualifiers
1..925
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMA534"
/tissue_type="mycelia"
/clone_lib="FVM"
/note="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: *Fusarium verticillioides*. Library
FVM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to Trizol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml Trizol. The cDNA was
directionally ligated into the pBluescript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

Query Match 10.8%; Score 312.2; DB 10; Length 925;
Best Local Similarity 59.9%; Pred. No. 8.4e-63;
Matches 548; Conservative 0; Mismatches 363; Indels 4; Gaps 2;

QY 1312 TTCTCCCAACAAATCCCTACCGCTCTAGCCAAATGGCGTTACCAACCATGTTTGGAGCG 1371
DB 10 TCTGCCCTCAGCAAGTACCCGAGCTCTTGATCTGTTAACCAACCATGTTGGCGGTG 69
QY 1372 GCACAGGTCTGTAGATGCGAATGCGACTACTATCACTCCGGGCAATGGAATCTTC 1431
DB 70 GTACCGGCCCAAGTCTCGAAGCAACGCAACTACTTGTAGCGCTGTGCTCATTCATG 129
QY 1432 ACCGATGTTGGCGCAGCAGAGAGTATCTTATGAATGTGGCTTTTGGGCAAGGCA 1491
DB 130 GTCAAAATGTTGAGCGGTGCGATCAGTTCCTTATCAATATTTGATTTACTGGCAAGGTA 189
|||||

```
QY 1492 ATAGCTCTAGCAAAAACAACTTGTAGAACAGTAGAAGCGGCGCGATGCTTTAAAT 1551
Db 190 ATGATAGTTCTCTGAGGGTCTGCGCATCAGGTCAGTCTGCTGTGGCTCAAGC 249
QY 1552 TGCAATGAAGACTGGGCGACAAACAAAGTGGATCGATCACTGCTGAGCGTGCAGATG 1611
Db 250 TTCAATGAGGACTGGGTTGCACTCTGCTGCTATGAGCTTGTCTAGTGTCTGTGATG 309
QY 1612 AATACGATGTGCAAGTTTGTATCCACCGCATACAGTCAATGAGGCGAGTTATGATG 1671
Db 310 AATTCCGATATTCAATGCTTATTACACTGACACGCTTAACGAGTCTGGCTTGTCTGAAT 369
QY 1672 ACACCCTAAATGCAATGAACGGGCGCGCATCCATGCTTACACATTTAGAGGCGGGTG 1731
Db 370 CTACGATCGTCTGCTTCAAGAACCGCAATTCATCTATCTATCAACAGAGGCTGACGAG 429
QY 1732 GAGGACACTCACTGATGTTATACCACTGGCAGCGAGTCAATATCTACCTCCCTCCA 1791
Db 430 GTGGCCATGCTCCGATATCATCTCGTGTGAGAGCATCAAAATGTTCTGCCATCATGA 489
QY 1792 CCACCCCTCACTATTCCCTATACCAATTAATACGGTTGAGAACACTTAGACATGCTCATGA 1851
Db 490 CCAACCTTACAAGACCACTTACACGCAATCTCTCGATGAGCATCTCGATATGCTTATGG 549
QY 1852 CATGCCACACTAGACAAACGATCCGCGAGGATTTACAATTTTCTCAAGCCGTATCC 1911
Db 550 TCTGCCATCACTTGTCCAAGAAATATCCAGAGGATGTAGCCCTTCGCGAGAGCGGTATTC 609
QY 1912 GCCCGGCTCTATCGCGCTGAAGATGTCCTCATGATATGCTGATGCGATGACAA 1971
Db 610 GTGCTGAACCAATGCTGCTGAGGATGTTATACGACAAAGGCGCTATCAGCATGATGA 669
QY 1972 GCTCGGATTCGCAAGCAATGGGCGTGCAGCGCAAGTGAATTCCTCGAATTTGGCAGACTG 2031
Db 670 GCTCTGACTCGCAGCTATGGCGCTTCCGAGAGGTCGTTTAAAGAACATGGAATCTG 729
QY 2032 CGGATAGAATAAAAAAGAAATTTGGTAGCTTCTGGAAGAT---GGCAAGATTAACGATA 2088
Db 730 CGCAATAAAATAAGGTGCAGAGGGGTTGTTGCGGAGGATGAGGCGACAGGGGCTGATA 789
QY 2089 ATTTCGCAATTAAGCGCTACATCTCCAAATACATATCAACCCGCTTTGACCCACGGCG 2148
Db 790 ATGCGCGTGTAAACGCTATGTCAGCAAGTATCTATTAAACCCAGCTATGCTCAGGGCT 849
QY 2149 TGACGAGTATATCGCTCTGTGGAAGAGGCAAGATCGCGCA-CTTGGTGGTGGGAAT 2207
Db 850 TTGACATGTCATTTGGAGCAATGAGGTGGAAAGTTTGTGATCTGTCTTTTGGAT 909
QY 2208 CTTGCCCTTTTGGC 2222
Db 910 CTTGCGTGGTTGGGC 924
```

```
RESULT 9
COL122599
LOCUS COL122599
DEFINITION GR_Eb04F01.f GR_Eb Gossypium raimondii cDNA clone GR_Eb04F01 5',
mRNA sequence.
ACCESSION COL122599
VERSION COL122599.1
KEYWORDS EST.
SOURCE COL122599.1 GI:48821286
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 872)
AUTHORS Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
```

Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 04 row: F column: 01.
Location/Qualifiers
1. 872
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb04F01"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

FEATURES

source

ORIGIN

```
Query Match 10.7%; Score 309; DB 8; Length 872;
Best Local Similarity 59.7%; Pred. No. 4.8e-62;
Matches 519; Conservative 0; Mismatches 350; Indels 0; Gaps 0;
QY 1260 GGGGAAGGTATGATTATTACCGCTGGGGGAATCGATTACACACCCACTTCTCTTCTCCA 1319
Db 2 GGAGAGAGTATGATCGTAACCTGACGGGCTATTGACTGCTGATGCTATTCATATGCTCT 61
QY 1320 CAACAAATCCCTACCGCTTAGCCAAATGGCGTTACAACCATGTTTGGAGCGGCAAGGT 1379
Db 62 CAGTTGGTTTCATGAAGCTATTATTCAGTGGCATCAACAATTTAGTTGGAGGTGGACCGGA 121
QY 1380 CTTGTAGATGCAAGAAATCGGACTACTATCATCTCCGGGCAATGGAATCTGCACCCATG 1439
Db 122 CCAGCTGAAGGAACACGTCGAACCTACTTGTACGCCAGCTCCGTCGCAATGAAATGATG 181
QY 1440 TTGGCGCAGCAGAGAGTATTCTATGATGTGGGCTTTTGGGCAAAAGGCAATAGCTCT 1499
Db 182 CTGAGTCCACTGATGACTTCCCTCTAAATTTGGCTTCACAGGAGGAAATGTTCT 241
QY 1500 AGCAAAAACAACTTTGTAGAACAGTAGAAGCGGCGCGGATTTGGTTTAAATTTGCATGA 1559
Db 242 AAACCTGAAGAAATACATGAATAATCAAAGCGGAGCAATGGACTGAAACTGCATGAG 301
QY 1560 GACTGGGCGCAACAACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATAGAT 1619
Db 302 GATTGGGGAACCTACACCTGCTGCAATAGACAGTTGTTGGCTGTTGCGAAGACTATATGAT 361
QY 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACCCCTA 1679
Db 362 ATCCAAAGTTAATATTCTACTGACACCTTTGAACGAATCTGGAATTTGGAAACAAACAT 421
QY 1680 AATGCAATGAACGGGCGGCCATCCATGCTACCACTTGAAGGAGCGGTTGAGGACAC 1739
Db 422 AATGCAATTTAAGGAAGAACTATTACACTTATCAAGTGAAGGTGCGGTGCTCAT 481
QY 1740 TCACCTGATTTATCAACCAATGGCAGCGAGCTCAATATTCTACCTCTCCACACCCCC 1799
Db 482 GCTCAGATATCATCAAAGTATGTGGCGTTAAACAGCTCTCCCTTCGTCAACAAACCCG 541
QY 1800 ACTATTCCTTATACCATTAATACGGTTGCAGAACCTTAGACATGCTCATGATGCCAC 1859
Db 542 ACTGCCCTTTATCTTCCAAATACTATAGATGAACATCTTGACATGCTGATGTTGCCAT 601
QY 1860 CACCTAGCAAAACGCAATCCGAGGATTTACAATTTCTCAAAGCCGTATCCGCCCGGC 1919
Db 602 CACCTCAGCAGGATATTCTGGAAGATGTAGCATTTGCAAGATCAAGGATAGGCGGGA 661
QY 1920 TCTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGATGCGGATGACAAAGCTCGAT 1979
```

Db	181	GAGATCGTCGAGGCGCGCATCGGGTCTCAAGTCCACGAGACTCGGGGACACACGCCG	240
Qy	1578	AGTGCAGATCGATCACTGCTTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCAC	1637
Db	241	TCAGCCATCATCAGCTGTCTTGAGCTCGGAGACAAAGTACGAGTGCAGGTCAATATTCAT	300
Qy	1638	ACCGATACAGTCAATGAGGCGAGTATTGTAGATGACACCTTAATGCAATGAACGGGCGC	1697
Db	301	ACGGATACGCTCAACGAGAGTGGATTGTTTGAGAGCACTATCAAGGCATTTCGGAAATCGC	360
Qy	1698	GCATTCATCGCTACCACTATGAGAGGAGCGGGTGGAGGACACTCACTGATGTTATTCACC	1757
Db	361	ACCAATTCACACTTATCACACCGAGGGTGC CGGTGGGGACACGCACTGATATCATCACC	420

[illegible]

Db	781	TCCAGTACACGATCAACCGGGCATACCGCATGGATGACCACTCGTTGGACAGTCC	840
Qy	2172	GAAGAGGGCAAGATCGCCGACTTTGGTGGTGTGGAATCCTGCTTTTTG	2220
Db	841	GCGGTCGGGGCGGCTCGGGACCTCGTCTCTGGAAGCCCGAGAACTTG	889
<p>RESULT 11</p> <p>CX676965</p> <p>LOCUS</p> <p>DEFINITION</p> <p>CX676965 748 bp mRNA linear EST 19-JAN-2005</p> <p>ydd29h04.y1 Sea urchin EST Lib1 Strongylocentrotus purpuratus cdna</p> <p>clone ydd29h04 5' similar to TR:Q9SR52 Q9SR52 PUTATIVE URASE. i,</p> <p>mRNA sequence.</p>			

CK676965
 CK676965.1 GI:57936920
 EST.
 STRONGYLOCENTROTUS PURPURATUS
 STRONGYLOCENTROTUS PURPURATUS
 Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinoidea; Euechinoidea; Echinacea; Echinoida;
 Strongylocentrotidae; Strongylocentrotus.
 1 (bases 1 to 748)
 COFFMAN, J.A., ROBERTSON, A.J., CLIFTON, S., PAPE, D., HILLIER, L.,
 MARTIN, J., WYLIE, T., DANTE, M., MEYER, R., THEISING, B., BOWERS, Y.,
 GIBBONS, M., RONKO, I., TEAGAREISHVILI, R., RITTER, E., KENNEDY, S. and
 WILSON, R.
 WASHU SEA URCHIN EST PROJECT
 UNPUBLISHED (2004)
 CONTACT: DR. JAMES A. COFFMAN
 WASHU SEA URCHIN EST PROJECT
 WASHINGTON UNIVERSITY SCHOOL OF MEDICINE
 4444 FOREST PARK PARKWAY, BOX 8501, ST. LOUIS, MO 63108
 TEL: 314 286 1900

Fax: 314 286 1810
Email: est@watson.wustl.edu
DNA sequencing by: Washington University Genome Sequencing Center
Seq primer: -28RppOT
High quality sequence stop: 680.
Location/Qualifiers
source
1..748
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="ydd29h04"
/lab_host="DH10B"
/clone_lib="Sea urchin EST Lib1"
/note="Vector: pCMVSPORT6.1 (Invitrogen); Site 1: NotI;
Site 2: SmaI; Arrayed normalized library of full-length
cDNAs representing blastula stage transcriptome of the sea
urchin Strongylocentrotus purpuratus, cloned into the
vector pCMVSPORT6.1 (Invitrogen)"

ORIGIN
Query Match 10.5%; Score 302.2; DB 9; Length 748;
Best Local Similarity 63.1%; Pred. No. 1.9e-60;
Matches 466; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
Qy 1648 TCAATGAGGAGGTATGTAGATGACACCCCTAAATGCAATGAACGGCGCGCCATCCATG 1707
Db |||||
Qy 1708 CCTACCATTTGAGGAGCGGTGAGGACACTCACCTGATGTATACCATGCGAGCG 1767
Db |||||
Qy 69 CCTATCATCTCGAGGAGTGCAGAGAGGCGACGCCCTGACATCATGAAGTAGTGGCG 128
Db |||||
Qy 1768 AGCTCAATATTCTACCCCTCTCCACCCGCCACTATTCCCTATACCAATTAAATACGGTTG 1827
Db |||||
Qy 129 TGCCCAATGTTACCTCGTCTACCAACCAACCGGCCCTTCACATGTAACCATCG 188
Db |||||
Qy 1828 CAGAACCTTAGACATGCTCATGATGCCACCATAGACAAAGCGATCCGAGGAT 1887
Db |||||
Qy 189 ATGAACACCTGGATATGCTATGCTGCCACCATCTTGAAGAATCTAAAGGAAGACG 248
Db |||||
Qy 1888 TACAATTTCTCAAGCGTATCCGCCGGCTCTATCGCGCTGAAGATGTCTCCATG 1947
Db |||||
Qy 249 TTGCTTTGACAGGTCTCGCATCAGACGCAACCAATGCTGCTGAAGATATCCTTCATG 308
Db |||||
Qy 1948 ATATGGGTGTGATCGGATGACAACTCGGATTCGCAAGCAATGGCGGTGAGCGCAAG 2007
Db |||||
Qy 309 ATCTGGAGCAATCAGTATTGGCTCGATTTCGAGGCCATGGGCGTGTGGTAGG 368
Db |||||
Qy 2008 TGATTCCTCGAACTTGGCAGACTCGGATAGAATAAATAAGAAATTTGGTAAGCTTCTG 2067
Db |||||
Qy 369 TGATCACCGGTACATGGCAGACGGCAGATAAGATGAAGATTTTCAGAGGCGAGCTGTG 428
Db |||||
Qy 2068 AAGATGGCAAGATAACGATAATTTCCGATTAAGCGGTATCTCCAAATACACTATCA 2127
Db |||||
Qy 429 AGSAAACGGGTGACAAACGACCAACCTTAAGCGTTACATTTCCCAAGTACACGGTCA 488
Db |||||
Qy 2128 ACCCGCTTTTCACCCAGCGGTGACGAGTATATCGGCTCTGTGGAAGAGGCGAAGTCG 2187
Db |||||
Qy 489 ACCCGCAATGCCCAGCGCATGCGCATCTCATTTGGCTCTGTACAGGTTGGTAATGG 548
Db |||||
Qy 2188 CCGACTTGGTGTGGTAATCTGCTTTTGGGTAAACCCCAAAATCGTGATCAAG 2247
Db |||||
Qy 549 CTGACCTGTTCTCTGGAATCGGCATTTCTCGAGACTTAACCTTGACCTCATCAAG 608
Db |||||
Qy 2248 CGGTATGGTGTCTTCTGGAATGGGGATTTAAACGGTCTGTGCGCCACTCCCAAC 2307
Db |||||
Qy 609 GAGGTATACATGCTTTGGGCACAAATGGGCGATGCCCAATGCATCCATCCGACCCCTGAAC 668
Db |||||
Qy 2308 CGGTATTATTCGCGAATGTTGGGCATCAGGCAAGCGGAATTTTCACACCCAGCATCA 2367
Db |||||
Qy 669 CTGTCAAGACGAGGAAATGTTGGTGTCTTACGGGAAGTCCATCGTGAGAACTCTGTCA 728
Db |||||
Qy 2368 CTTTGTGTTTCCAAAGTCGC 2386

Db 729 TCTTTGTGTCCTCAAGCAGC 747
|||||
RESULT 12
LOCUS DR637386 889 bp mRNA linear EST 11-JUL-2005
DEFINITION EST1028011 FvM Gibberella moniliformis cDNA clone FVMAV95, mRNA
sequence.
ACCESSION DR637386
VERSION DR637386.1 GI:70712220
KEYWORDS EST
SOURCE Gibberella moniliformis
ORGANISM Gibberella moniliformis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 889)
AUTHORS Brown,D.W., Cheung,F., Proctor,R.H., Butchko,A.E., Zheng,L., Lee,Y.,
Utterback,T., Smith,S., Feidblyum,T., Glenn,A.E., Plattner,R.D.,
Kendra,D.F., Town,C.D. and Whitelaw,C.A.
Comparative analysis of 87,000 expressed sequence tags from the
fungus-in-producing fungus Fusarium verticillioides
Fungal Genet. Biol. 42 (10), 848-861 (2005)
16099185
Contact: Brown, D.W.
USDA/ARS/NCAR
USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncar.usda.gov
TIGR sequence name: FVMAV95TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
FEATURES
Location/Qualifiers
1..889
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAV95"
/tissue_type="mycelia"
/clone_lib="FvM"
/note="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYAM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to Trizol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml Trizol. The cDNA was
directionally ligated into the pBlueScript II SK(+) XR
vector (CDNA Synthesis Kit; Stratagene)."
ORIGIN
Query Match 10.4%; Score 300.2; DB 10; Length 889;
Best Local Similarity 59.5%; Pred. No. 5.9e-60;
Matches 524; Conservative 0; Mismatches 353; Indels 3; Gaps 1;
Qy 1312 TTTCTCCACAAATTCCTACCGCTCTAGCCAAATGGCGTTACACCATGTTGGAGCG 1371
Db |||||
Qy 10 TCTGCCCTCAGCAAGTACCCGAAGCTCTTGCACTGTGTAAACCACTGCTTGGCGGT 69
Db |||||
Qy 1372 GCACAGGTCTCTAGATGGCAGAAATGCGACTACTATCACTCCGGGCAATGGAACCTGC 1431
Db |||||
Qy 70 GTACCGGCCCAAGTCTGGAACGAACGCAACTACTGTAGCGCTGTGTCTATCATGC 129
Db |||||
Qy 1432 ACCGCATGTTGCGCGCAGCAGAAAGATTTCTATGAATGTGGCTTTTGGGCAAGGCA 1491
Db |||||
Qy 130 GTCAATGTTGCGCGGTGCGATCAGCTTCCTATCAATATTTGTTACTTGGCAAGGTA 189
Db |||||
Qy 1492 ATAGCTCTAGCAAAAAAACAATTTGTAGAACAGTAGAACGGCGCGATGTTTAAAT 1551
Db |||||
Qy 190 ATGATAGTTCTCCTGAGGCTCTGCGCGATCAGTCAATGCTGTGTCTTGTGCTCAAGC 249
Db |||||


```

Db      663  GCGGTATTGCTGCTGAACCAATTGCTGCTGAGGATGATTATACAGCAAAAGCGCGTATCA 722
Qy      1363  CGATGACAAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGCGGAAGTGAATTCCTCGAACTT 2022
Db      723  GCATGATGAGCTCTGACTCGCAGGCTATGGCGGTGCGGAGAGGTGCTTTTAAGAACAAT 782
Qy      2023  GGCAGACTGGCGAT 2036
Db      783  GGAATACTGCGCAT 796

RESULT 14
DR635599
LOCUS      EST1026224 FvM Gibberella moniliformis cDNA clone FVMAJ96, mRNA
DEFINITION      849 bp  mRNA  linear  EST 11-JUL-2005
ACCESSION      DR635599
VERSION        DR635599.1  GI:70710433
KEYWORDS       Gibberella moniliformis
SOURCE         Gibberella moniliformis
ORGANISM       Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
                Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE      1 (bases 1 to 849)
AUTHORS        Brown,D.W., Cheung,F., Proctor,R.H., Butchko,A.E., Zheng,L, Lee,Y.,
                Uterback,T., Smith,S., Feldblyum,T., Glenn,A.E., Plattner,R.D.,
                Kendra,D.F., Town,C.D. and Whitelaw,C.A.
TITLE          Comparative analysis of 87,000 expressed sequence tags from the
                fumonisin-producing fungus Fusarium verticillioides
JOURNAL        Fungal Genet. Biol. 42 (10), 848-861 (2005)
PUBMED         16099185
COMMENT        Contact: Brown, D.W.
                USDA/ARS/NCAUR
                1815 N. University St, Peoria, IL 61604, USA
                Tel: 309 681 6230
                Fax: 309 681 6689
                Email: brown@ncaur.usda.gov
                TIGR sequence name: FVMAJ96TH
                Seq primer: AAT TAA CCC TCA CTA AAG GG.
                Location/Qualifiers
                1..849
                /organism="Gibberella moniliformis"
                /mol_type="mRNA"
                /strain="m3125"
                /db_xref="taxon:117187"
                /clone="FVMAJ96"
                /tissue_type="mycelia"
                /clone_lib="FvM"
                /note="Vector: pBlueScript II SK(+); Site 1: EcoRI;
                Site 2: XhoI; anamorph: Fusarium verticillioides. Library
                FvM was prepared from pooled RNA obtained from a 48-hour
                and a 72 hour, liquid GYM culture from strain M-3125.
                Cultures were vacuum filtered and the mycelial mats were
                frozen in liquid nitrogen, ground to a powder, and then
                added to Trizol Reagent (Invitrogen, Carlsbad CA) at
                approximately 1 g mycelia per 10 ml Trizol. The cDNA was
                directionally ligated into the pBlueScript II SK(+) XR
                vector (cDNA Synthesis Kit; Stratagene)."
```

ORIGIN

```

Query Match      10.2%; Score 293; DB 10; Length 849;
Best Local Similarity 60.2%; Pred No. 3e-58;
Matches 503; Conservative 0; Mismatches 330; Indels 3; Gaps 1;

Qy      1317  CCACAACAATTCCTACCGCTCTAGCCAAATGGCGTTACAAACCATGTTGGAGCGGCACA 1376
Db      3    CCTCAGCAAGTACCGGAAGCTCTTGCATCTGCTCTAAACCAACCATGCTTGGCGGTGTACC 62
Qy      1377  GGTCTGTAGATGGCAAGATCGGACTACTATCTCCGGGCAATGGAACCTTGACCGC 1436
Db      63  GGCCCAAGTGTGGAAACGAACCAACTACTTGTACGCGCTGTGCTCAATTACATGCGTCAA 122
```

```

Qy      1437  ATGTTGGCGGAGCAGAGAGTATTCTATGAATGTGGGCTTTTGGGCAAGCAATAGC 1496
Db      123  ATGTTGCAAGGCTGGGATCAGCTTCTCATCAATATTGGTATTACTGSCAAAGTAATGAT 182
Qy      1497  TCTACAAAAAACAACCTTGTAGAACAAAGTAGAAGCGGCGCGATTGCTTTAAATTGTCAT 1556
Db      183  AGTTCTCTGAGGGTCTGGCGGATCAGGTCAATGCTGGTGTCTCAGTGTCTGTGATGAATTC 242
Qy      1557  GAAAGACTGGGGCACAACCAAGTGGCATCGATCACTGCTTGAGCGGTGGCAGATGAATAC 1616
Db      243  GAGGACTGGGGTTGCACCTCTGCTGCTATTGACGCTTGTCTCAGTGTCTGTGATGAATTC 302
Qy      1617  GATGTGAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGGTTATGTAGATGACACC 1676
Db      303  GATATTCAATGTCTTATTTCACACTGACACGCTTAACGAGTGTGGCTTTGTGCAATCTACG 362
Qy      1677  CTAATGAATGAACGCGGCGCCATCCATGCTACCACTGAGGAGCGGGTGGAGGA 1736
Db      363  ATCGCTGCTTTCAAGAACCGCAATTCATCTTATCACAGAGGTTGACAGAGGTGGC 422
Qy      1737  CACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCCACCAACC 1796
Db      423  CATGCTCCGGATATCATCTCCGTTGGTAGAGCATCAAAATGTTCTGCCATCATCGACCAAC 482
Qy      1797  CCCATATTCCTATACCAATTAATAGGTTGAGAACTTACAGATGCTATGATGACATGC 1856
Db      483  CCTACAAGACCAATTCACACGCAATATCTCTCGATGAGCATCTCGATATGCTTATGGTTCG 542
Qy      1857  CACCACTTAGAACAACGCGATCCGCGAGGATTACAAATTTTCTCAAGCCGTATCCGCCCC 1916
Db      543  CATCACTTGTCCAAGAAATCCAGAGGATGTAGCCTTCGCGAGAGCCGTATTCGTGCT 602
Qy      1917  GGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCG 1976
Db      603  GAAACCATGTGCTGAGGATGTATTACACGACAAAGCGCTATCAGCATGATGAGCTCT 662
Qy      1977  GATTCGCAAGCAATGGGCGTGCAGCGAAGTGAATTCCTGAACTTGGCAGACTGCGGAT 2036
Db      663  GACTCGCAGGCTATGGCGGTTCGCGAGAGGTGCTTTTAAAGAACATGGAATATCGCGCAT 722
Qy      2037  AGAATAAAAAGAATTTGTTAAGCTTCTCTGAGAT---GGCAAGATAACGATAATTTC 2093
Db      723  AAAATAAAGTGCAGAGGGTTGGTTGCCGAGAGTGAAGGGGCACAGGGCTGATAATGCG 782
Qy      2094  CGCATTAAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACCCACGCGGT 2149
Db      783  CGTGTAAAACGCTATGCTCAGCAAGTACTATTAAACCCAGCTATTGCTCAGGCGTT 838
```

RESULT 15

```

AZ935182/c
LOCUS      784 bp  DNA  linear  GSS 24-APR-2001
DEFINITION      BJ_Ba003020f B. japonicum BAC library Bradyrhizobium japonicum
                genomic, genomic survey sequence.
ACCESSION      AZ935182
VERSION        AZ935182.1  GI:13777494
KEYWORDS       GSS.
SOURCE         Bradyrhizobium japonicum
ORGANISM       Bradyrhizobium japonicum
```

```

Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Bradyrhizobium.
1 (bases 1 to 784)
Tomkins, J. P., Wood, T. C., Stacey, M. G., Loh, J. T., Judd, A.,
Goicoechea, J. L., Stacey, G., Sadowsky, M. J. and Wing, R. A.
A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
genome
Genome Res. 11 (8), 1434-1440 (2001)
11483585
Contact: Wing RA
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
```

REFERENCE

11483585

AUTHORS

11483585

TITLE

11483585

JOURNAL

11483585

PUBMED

11483585

COMMENT

11483585

Fax: 864 656 4293
 Email: rwing@clmson.edu
 Class: BAC ends
 High quality sequence stop: 719.
 Location/Qualifiers
 1..784
 /organism="Bradyrhizobium japonicum"
 /mol_type="genomic DNA"
 /strain="USD110"
 /db_xref="taxon:375"
 /lab_host="E. coli"
 /clone_lib="B. japonicum BAC library"
 /note="Vector: pindigo536; Site_1: HindIII"

FEATURES
 source

ORIGIN

Query Match	9.9%;	Score 284;	DB 11;	Length 784;
Best Local Similarity	64.0%;	Pred. No. 4.1e-56;		
Matches 458;	Conservative 0;	Mismatches 253;	Indels 5;	Gaps 3;

Qy	1468	ATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAACAACTTTGTAGAACAAAGTAG	1527
Db	715	ATCTCGGCAATTCGGGCAAGGCAACGCTCGCGCCCGCGCTGTGTCGAGATCATCA	656
Qy	1528	AAGCGGGCGGATGCTTTAAATTGCATGAGACTGGGGCACACCAAGTGGATCG	1587
Db	655	AGNCGGGCGCATGCGCGCTGAAGCTGCACGAGGATTGGGGCAC-ACGCGCGCGCGATCG	597
Qy	1588	ATCACTGCTTGAGCGTGGCAGATGAATACGATGTCAAGTTTGTATCCACACCGATACAG	1647
Db	596	ACAACTGCTGTGGTGGCCGACGATACGATCCAGTCTATGATCCACCCGATACGC	537
Qy	1648	TCAATGAGGCAAGTTATAGATGACACCCCTAAATGCAATGAACGGGCGCGCCATCCATG	1707
Db	536	TGACGAATCCGGCTTCGTCGAGGATACGATCAAGGCGTTCAAGGGCGCACCATCCACG	477
Qy	1708	CCTACCACATTTAGGGAGCGGTGGAGACACTCACTGATGTTATACCATGGCAGGG	1767
Db	476	CTTTCCACACCGAGGGCGCGCGCGGTACGCCCGCGACATCAAGGTTCGCAAGGC	417
Qy	1768	AGCTCAATATTCTACCTCTCCACCCACCTATTTCCCTATACCATTAATACGGTTG	1827
Db	416	TGAAGACGTCGTGCGGTCATCGAACACCCGCGCGCCCTTACCCGCAACCATCG	357
Qy	1828	CAGAACACTTAGACATGCTCATGATCATGCCACCACTAGACAAAGCATCCGCGAGGATT	1887
Db	356	ACGACATCTCGACATGCTGATGTTGTCACCACTCGATCCCTCGATCGCGAAGATC	297
Qy	1888	TACAAATTTCTCAAGCGGATACCGCCCGCGCTTATCGCGGCTGAAGATGTGTCATG	1947
Db	296	TTGGCTTCGCCGAAGCGGATATCGCAAGGAGACCATCGCGCGCGAGACATCCTGCACG	237
Qy	1948	ATATGGGTGTCATCGCGATGACAGCTCGGATTCGACAGCAATGGG-CGCTCAGGGCGAA	2006
Db	236	ATCTCGGCGCGCTCTCGATGATGTCCTCGNACTCCAGGCCATGGGCGCGCTCGGCGAA	177
Qy	2007	GTGATTCCTCGAACTTGGCAGACTGGGATGAAGATAAAGAAATTTGGTAAGCTTCC-	2065
Db	176	GTATCATCCGGACTTGGCAGACCGCGCAAGATGAAGAGCAGCGCGGATCGCTGCGG	117
Qy	2066	--TGAAGATGCCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACT	2123
Db	116	CAGGACAAAGGCAAGGACAAAGCAATTTCCGCGTCAAGCGCTACATCGCCAAATACAG	57
Qy	2124	ATCAACCCCGCTTTGACCCCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGG	2179
Db	56	ATCAACNCCGGATCGCGCACGGGTGTCGAAGCTGATCGGTTTCGTTGGAAGAGGG	1

Search completed: August 10, 2006, 05:05:16
 Job time : 12173 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:20:07 ; Search time 195 Seconds
(without alignments)
529.902 Million cell updates/sec

Title: US-09-904-994b-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLYYAGEVA.....ADBRHKHAKDKXSHGFIK 226

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_8.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*
- 10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1181	100.0	226	5	ADJ58238 Urease su
2	1178	99.7	226	5	Adj58250 UreaseXY
3	1174	99.4	226	5	ADJ58241 UreaseXY
4	1174	99.4	226	5	ADJ58247 UreaseXY
5	1170	99.1	226	5	ADJ58244 UreaseXY
6	665	56.3	234	8	ADQ37849 H. bizzoz
7	619	52.4	238	4	Aau35875 Helicobac
8	616	52.2	238	2	AAR04579 Part of p
9	616	52.2	238	2	Aaw07193 H. pylori
10	616	52.2	238	4	Aau35693 Helicobac
11	616	52.2	238	5	Aam49644 Urease A
12	616	52.2	245	4	AEB55115 Helicobac
13	613	51.9	238	8	ADM28643 Helicobac
14	609	51.6	238	8	ADM28643 Helicobac
15	608	51.5	238	2	AAR67375 H. pylori
16	602	51.0	238	5	Aam49645 Salmonsell
17	601	50.9	238	2	AAR12515 A subunit
18	579	49.0	228	3	AAB52559 Helicobac
19	579	49.0	228	5	ABU1893 Helicobac
20	578	48.9	237	2	AAR74336 Helicobac
21	578	48.9	237	2	Aaw06729 H. felis
22	578	48.9	806	2	AAR67371 Urease A
23	568	48.1	222	5	Aam49647 Gerbil ur

24	565	47.8	219	5	AAM49646	Aam49646 Murine ur
25	565	47.8	224	5	AAM49648	Aam49648 Murine ur
26	530	44.9	811	6	ABU42650	ABU42650 Protein e
27	516	43.7	207	6	ABU42523	ABU42523 Protein e
28	509	43.1	226	4	AAU33952	AAU33952 Staphyloc
29	505	42.8	779	6	ABU41210	ABU41210 Protein e
30	493	41.7	207	6	ABU17114	ABU17114 Protein e
31	492	41.7	207	6	ABU20285	ABU20285 Protein e
32	490	41.5	261	6	ABU27583	ABU27583 Protein e
33	487	41.2	209	6	ABU40327	ABU40327 Protein e
34	478	40.5	1216	2	AAR67376	AAR67376 P. mirabi
35	477	40.4	209	2	AAR67376	AAR67376 P. mirabi
36	467	39.5	228	6	ABU22153	ABU22153 Protein e
37	462	39.1	840	8	ADJ66032	ADJ66032 Coil form
38	462	39.1	840	9	AEA54210	AEA54210 C. ensifo
39	459	38.9	161	5	ABU50981	ABU50981 Helicobac
40	449	38.0	270	2	AAR67377	AAR67377 Canavalia
41	382	32.3	837	8	ADG65522	ADG65522 C. posada
42	291	24.6	121	6	ABU21848	ABU21848 Protein e
43	287	24.3	100	2	AAW37775	AAW37775 Klebsiell
44	287	24.3	100	3	AAy81823	AAy81823 Klebsiell
45	287	24.3	100	4	AAU36152	AAU36152 Klebsiell

ALIGNMENTS

RESULT 1
ADJ58238
ID ADJ58238 standard; protein; 226 AA.
XX
AC ADJ58238;
XX
DT 06-MAY-2004 (first entry)
XX
DE Urease subunit polypeptide X.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
PN EPI1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
DR N-PSDB; ADJ58237.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Claim 8; SEQ ID NO 2; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an urease
CC X subunit polypeptide of the invention.
XX
SQ Sequence 226 AA;

```
Query Match      100.0%; Score 1181; DB 5; Length 226;
Best Local Similarity 100.0%; Pred. No. 2.1e-45;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKRKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||
Db 1 VKLTPKEQKFLYYAGEVARKKRKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||

QY 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDEHFHFKAGEVKFGCDKDIE 120
   |||
Db 1 VKLTPKEQKFLYYAGEVARKKRKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||

QY 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDEHFHFKAGEVKFGCDKDIE 120
   |||
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDEHFHFKAGEVKFGCDKDIE 120
   |||

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFANKALKFDREKAYGKRDLIPSGNTLRIGAG 180
   |||
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDEHFHFKAGEVKFGCDKDIE 120
   |||

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFANKALKFDREKAYGKRDLIPSGNTLRIGAG 180
   |||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFANKALKFDREKAYGKRDLIPSGNTLRIGAG 180
   |||

QY 181 QTRKVQLIPLGGSKKVIWMNGLVNNIADERHKKHAKLDKAKSHGFIK 226
   |||
Db 181 QTRKVQLIPLGGSKKVIWMNGLVNNIADERHKKHAKLDKAKSHGFIK 226
   |||

RESULT 2
ADJ58250
ID ADJ58250 standard; protein; 226 AA.
XX
XX ADJ58250;
XX
XX 06-MAY-2004 (first entry)
XX
XX UreaseXY subunit #7.
XX
XX immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
XX Helicobacter felis.
XX
XX EP1176192-A2.
XX
XX 30-JAN-2002.
XX
XX 11-JUL-2001; 2001EP-00202666.
XX
XX 17-JUL-2000; 2000EP-00202565.
XX
XX (ALKU ) AKZO NOBEL NV.
XX
XX Kusters JG, Cattoli G;
XX
XX WPI; 2002-124384/17.
XX
XX N-PSDB; ADJ58249.
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
XX Disclosure; SEQ ID NO 14; 76pp; English.
XX
XX The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
XX
XX Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an
ureaseXY subunit of the invention.
XX
XX Sequence 226 AA;

Query Match      99.7%; Score 1178; DB 5; Length 226;
Best Local Similarity 99.6%; Pred. No. 2.9e-45;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKRKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||
Db 1 VKLTPKEQKFLYYAGEVARKKRKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||

QY 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDEHFHFKAGEVKFGCDKDIE 120
   |||
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDEHFHFKAGEVKFGCDKDIE 120
   |||

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFANKALKFDREKAYGKRDLIPSGNTLRIGAG 180
   |||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHPFANKALKFDREKAYGKRDLIPSGNTLRIGAG 180
   |||

QY 181 QTRKVQLIPLGGSKKVIWMNGLVNNIADERHKKHAKLDKAKSHGFIK 226
   |||
Db 181 QTRKVQLIPLGGSKKVIWMNGLVNNIADERHKKHAKLDKAKSHGFIK 226
   |||

RESULT 3
ADJ58241
ID ADJ58241 standard; protein; 226 AA.
XX
XX ADJ58241;
XX
XX 06-MAY-2004 (first entry)
XX
XX UreaseXY subunit #1.
XX
XX immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
XX Helicobacter felis.
XX
XX EP1176192-A2.
XX
XX 30-JAN-2002.
XX
XX 11-JUL-2001; 2001EP-00202666.
XX
XX 17-JUL-2000; 2000EP-00202565.
XX
XX (ALKU ) AKZO NOBEL NV.
XX
XX Kusters JG, Cattoli G;
XX
XX WPI; 2002-124384/17.
XX
XX N-PSDB; ADJ58240.
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
XX Disclosure; SEQ ID NO 5; 76pp; English.
XX
XX The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
XX
XX Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an
ureaseXY subunit of the invention.
XX
XX Sequence 226 AA;

Query Match      99.4%; Score 1174; DB 5; Length 226;
Best Local Similarity 99.1%; Pred. No. 4.5e-45;
Matches 224; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKRKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||
Db 1 VKLTPKEQKFLYYAGEVARKKRKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||

QY 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDEHFHFKAGEVKFGCDKDIE 120
   |||
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDEHFHFKAGEVKFGCDKDIE 120
   |||
```

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVI GMGNLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMGNLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 4
ADJ58247
ID ADJ58247 standard; protein; 226 AA.
XX
XX AC ADJ58247;
XX 06-MAY-2004 (first entry)
XX UreaseXY subunit #5.
XX immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX OS Helicobacter felis.
XX PN EPI1176192-A2.
XX 30-JAN-2002.
XX 11-JUL-2001; 2001EP-00202666.
XX 17-JUL-2000; 2000EP-00202565.
XX (ALKU) AKZO NOBEL NV.
XX Kusters JG, Cattoli G;
XX WPI; 2002-124384/17.
XX N-PSDB; ADJ58246.
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX Disclosure; SEQ ID NO 11; 76pp; English.

XX The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.

XX Sequence 226 AA;
QY Query Match 99.4%; Score 1174; DB 5; Length 226;
Best Local Similarity 99.1%; Pred. No. 4.5e-45;
Matches 224; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 1 VKLTPKEQEKFLYYAGVARRKKAEGKLKQNPPEAIYISAHIMDEARRGKKTVAQLMEE 60
1 VKLTPKEQEKFLYYAGVARRKKAEGKLKQNPPEAIYISAHIMDEARRGKKTVAELMEE 60
QY 61 CMHFLKKDEVMGPNVGDILGVEATFPDGTGLVTNNWPIEPDHFPAKEVKGCDKDIE 120
Db 61 CMHFLKKDEVMGPNVGDILGVEATFPDGTGLVTNNWPIEPDHFPAKEVKGCDKDIE 120
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVI GMGNLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMGNLVNNIADERHKHKKALDKAKSHGFIK 226

Db 181 QTRKVQLIPLGGSKKVI GMGNLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 5
ADJ58244
ID ADJ58244 standard; protein; 226 AA.
XX
XX AC ADJ58244;
XX 06-MAY-2004 (first entry)
XX UreaseXY subunit #3.
XX immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX OS Helicobacter felis.
XX PN EPI1176192-A2.
XX 30-JAN-2002.
XX 11-JUL-2001; 2001EP-00202666.
XX 17-JUL-2000; 2000EP-00202565.
XX (ALKU) AKZO NOBEL NV.
XX Kusters JG, Cattoli G;
XX WPI; 2002-124384/17.
XX N-PSDB; ADJ58243.
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX Disclosure; SEQ ID NO 8; 76pp; English.

XX The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.

XX Sequence 226 AA;
QY Query Match 99.1%; Score 1170; DB 5; Length 226;
Best Local Similarity 98.7%; Pred. No. 6.8e-45;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 1 VKLTPKEQEKFLYYAGVARRKKAEGKLKQNPPEAIYISAHIMDEARRGKKTVAQLMEE 60
1 VKLTPKEQEKFLYYAGVARRKKAEGKLKQNPPEAIYISAHIMDEARRGKKTVAELMEE 60
QY 61 CMHFLKKDEVMGPNVGDILGVEATFPDGTGLVTNNWPIEPDHFPAKEVKGCDKDIE 120
Db 61 CMHFLKKDEVMGPNVGDILGVEATFPDGTGLVTNNWPIEPDHFPAKEVKGCDKDIE 120
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVI GMGNLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMGNLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 6
ADQ37849
ID ADQ37849 standard; protein; 234 AA.

XX AC ADQ37849;
 XX DT 07-OCT-2004 (first entry)
 XX DE H. bizzozeronii urea polypeptide.
 XX KW Urease; urease gene cluster; urease structural gene;
 XX KW urease accessory gene; ureA; Helicobacter bizzozeronii infection;
 XX KW antibacterial; enzyme.
 XX OS Helicobacter bizzozeronii.
 XX PN US2004142343-A1.
 XX PD 22-JUL-2004.
 XX PF 12-AUG-2003; 2003US-00639273.
 XX PR 16-AUG-2002; 2002US-0404337P.
 XX PA (CHAN/) CHANG Y.
 XX PA (SIMP/) SIMPSON K W.
 XX PA (ZHUJ/) ZHU J.
 XX PI Chang Y, Simpson KW, Zhu J;
 XX WPI; 2004-533502/51.
 XX DR N-PSDB; ADQ37848.
 XX Novel isolated nucleic acid molecule having urease gene cluster, and
 PT conferring on Helicobacter bizzozeronii ability to produce urease, useful
 PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
 XX Claim 23; SEQ ID NO 3; 40pp; English.
 XX The invention relates to an isolated nucleic acid molecule conferring on
 CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
 CC acid molecule is a urease gene cluster comprising at least one urease
 CC structural gene and at least one urease accessory gene. The nucleic acid
 CC molecule is chosen from ureA, ureB, ureE, ureF, ureG, ureH and ureI. The
 CC invention also relates to an isolated protein encoded by the nucleic
 CC acid, a vaccine for preventing onset of disease in mammals infected by H.
 CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
 CC antibody or its binding portion raised against the nucleic acid. The
 CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
 CC against onset of disease caused by infection of H. bizzozeronii, which
 CC involves administering the sequences. The sequences are useful for
 CC detecting H. bizzozeronii in a sample of tissue or body fluids which
 CC or providing a nucleotide sequence as an antigen, providing an antibody,
 CC or providing a nucleotide sequence as a probe in a nucleic acid
 CC hybridisation assay, contacting the sample with the antigen or the probe,
 CC and detecting any reaction which indicates that H. bizzozeronii is
 CC present in the sample. This sequence represents the H. bizzozeronii ureA
 CC polypeptide.
 XX Sequence 234 AA;
 SQ Query Match 56.3%; Score 665; DB 8; Length 234;
 Best Local Similarity 57.3%; Pred. No. 1.2e-21;
 Matches 130; Conservative 39; Mismatches 52; Indels 6; Gaps 3;
 QY 1 VKLTPKEDEKFLLYVAGSVARAKRAGKLNQPEAIVISAHIMDEARRGKKTVAQLMEE 60
 DB 1 MKLTPKELDKLMLHVLGAKLRANGVKLNTTEVALISAHVMEARAGKKSVDLMOQE 60
 QY 61 CMHFLKKDEVNPGVGNVPDLGVZATFPDGTGKLVTVNWPIDPEHFKV--KFGCDK 117
 DB 61 GRTLLKADVVMPGVAHMEHVEGIEANFPDGTGKLVTIHTPVEDGGHKLAPGEVILK--NE 117
 QY 118 DIENAGKEVTELVNTPGKSLHVGSHFFPEANKALKFDEKAYGKRDLIPSGNTLRI 177
 DB 118 DITLGNAGQATTLVHNKGRDPVQVQSGSHFFPEVNVKLLFDEKAYGKRDLIASGTAVRF 177

QY 178 GAGQTRKVLQIPLGGSKKVGNGLVNINNIADRRHKKHAKLDKAKSHGF 224
 DB 178 EPGEKKTVELIQIGGNQRIYGFNSLVDRQADTDGKKLAKRAKERGF 224
 RESULT 7
 AAU35875
 ID AAU35875 standard; protein; 238 AA.
 XX AC AAU35875;
 XX DT 14-FEB-2002 (first entry)
 XX DE Helicobacter pylori cellular proliferation protein #188.
 XX KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX OS Helicobacter pylori.
 XX PN WO200170955-A2.
 XX PD 27-SEP-2001.
 XX PF 21-MAR-2001; 2001WO-US009180.
 XX PR 21-MAR-2000; 2000US-0191078P.
 XX PR 23-MAY-2000; 2000US-0206848P.
 XX PR 26-MAY-2000; 2000US-0207272P.
 XX PR 23-OCT-2000; 2000US-0245278P.
 XX PR 27-NOV-2000; 2000US-0253625P.
 XX PR 22-DEC-2000; 2000US-0257931P.
 XX PR 16-FEB-2001; 2001US-0269308P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI; 2001-611495/70.
 XX DR N-PSDB; AAS53734.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX Example 3; SEQ ID NO 11468; 511pp; English.
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 238 AA;
 SQ Query Match 52.4%; Score 619; DB 4; Length 238;
 Best Local Similarity 53.1%; Pred. No. 1.6e-19;
 Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKELLYYAGEVARKKAEGLKLNQPEAIAIVISAHIMDEARRGKKTVAQLMEE 60
 Db 1 MKLTPKELDKMLHYAGELAKRKKGKLNVEAVALISAHIMEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVPNGVMVDPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
 Db 61 GRTLKPDVDMGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLIRIGAG 180
 Db 120 INEGKKAHSVKNVGNDRPVGIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRPEPG 179
 QY 181 QTRKVLQILPGLGSKKVGIMGNGLVNNIADERHKHAKLDKAKSHGF 224
 Db 180 EEKVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 8
 AAR04579
 ID AAR04579 standard; protein; 238 AA.
 XX
 AC AAR04579;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 21-SEP-1990 (first entry)
 XX
 DE Part of protein with urease activity.
 XX
 KW Urease; probe.
 XX
 OS Helicobacter pylori.
 XX
 PN WO9004030-A.
 XX
 PD 19-APR-1990.
 XX
 PF 06-OCT-1988; 88FR-00013135.
 XX
 PR 06-OCT-1988; 88FR-00013135.
 XX
 PA (INSP) INST PASTEUR.
 PA (INRM) INSERM INST NAT SANTE & RECH MED.
 XX
 PI Labigne A;
 XX
 DR WPI; 1990-147844/19.
 DR N-PSDB; AAQ04328.
 XX
 PT New nucleotide sequences encoding Campylobacter pylori-ureaseants - and
 PT derived vectors, transformants, protein, antibodies and probes, useful in
 PT diagnosis, treatment and prevention of infections.
 XX
 PS Claim 11; Page 34; 47pp; French.
 XX
 CC The protein can be used for the prodn. of antibodies and to prepare
 CC vaccines for the prevention /treatment of C. pylori infections. See also
 CC AAR04578-81. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27
 CC -AUG-2003 to correct OS field.)
 XX
 SQ Sequence 238 AA;

Query Match 52.2%; Score 616; DB 2; Length 238;
 Best Local Similarity 52.7%; Pred. No. 2.2e-19;
 Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKELLYYAGEVARKKAEGLKLNQPEAIAIVISAHIMDEARRGKKTVAQLMEE 60
 Db 1 MKLTPKELDKMLHYAGELAKRKKGKLNVEAVALISAHIMEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVPNGVMVDPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
 Db 61 GRTLKPDVDMGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLIRIGAG 180
 Db 120 INEGKKAHSVKNVGNDRPVGIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRPEPG 179
 QY 181 QTRKVLQILPGLGSKKVGIMGNGLVNNIADERHKHAKLDKAKSHGF 224
 Db 180 EEKVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223

RESULT 9
 AAW07193
 ID AAW07193 standard; protein; 238 AA.
 XX
 AC AAW07193;
 XX
 DT 16-OCT-2003 (revised)
 DT 11-PEB-1997 (first entry)
 XX
 DE H. pylori urease A subunit.
 XX
 KW Urease; ureA gene; ureB gene; vaccine.
 XX
 OS Helicobacter pylori; strain CPM630.
 XX
 PN WO9633732-A1.
 XX
 PD 31-OCT-1996.
 XX
 PF 25-APR-1996; 96WO-US005800.
 XX
 PR 28-APR-1995; 95US-00431041.
 PR 06-DEC-1995; 95US-00568122.
 XX
 PA (ORAV-) ORAVAX INC.
 XX
 PI Lee CK, Monath TP, Ackerman SK, Thomas WD, Soman G, Kleanthous H;
 PI Weltzin RA, Pappo J, Ermak T, Guirakhoo F, Bhagat H, Sussman I;
 XX
 DR WPI; 1996-497373/49.
 DR N-PSDB; AAT44351.
 XX
 PT Vaccine for inducing mucosal response to Helicobacter - contg. multimeric
 PT urease complex and pref. an antibiotic, anti-secretory agent or bismuth
 PT salt.
 XX
 PS Disclosure; Page 70-71; 98pp; English.
 XX
 CC Urease A (AAW07193) and B (AAW07194) subunits of Helicobacter pylori are
 CC encoded by the ureA + ureB gene locus of clinical isolate CPM630.
 CC Vectors, e.g. pORV214 (see also AAT44351), carrying the gene locus can be
 CC used to transform host cells for the large-scale prodn. of recombinant,
 CC enzymatically inactive, multimeric urease. The urease complex induces a
 CC mucosal immune response that can treat or prevent Helicobacter, esp. H.
 CC pylori, gastrointestinal infection. (Updated on 16-OCT-2003 to standardise
 CC OS field)
 XX
 SQ Sequence 238 AA;

Query Match 52.2%; Score 616; DB 2; Length 238;
 Best Local Similarity 52.7%; Pred. No. 2.2e-19;
 Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKELLYYAGEVARKKAEGLKLNQPEAIAIVISAHIMDEARRGKKTVAQLMEE 60
 Db 1 MKLTPKELDKMLHYAGELAKRKKGKLNVEAVALISAHIMEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVPNGVMVDPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
 Db 61 GRTLKPDVDMGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLIRIGAG 180
 Db 120 INEGKKAHSVKNVGNDRPVGIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRPEPG 179

Db 120 INGGKAVSVKVNVDGPRPVOIGSHHFFVNVNRCDFDREKTFGKRLDIAGSTAVRFEG 170
 QY 181 QTRKVLPLGGSKVKTCMGNLVNNIADERHKHAKALDKAKSHGF 224
 Db 180 EKSVELIDIGGNRIFGFGNALVDQADNESKKTALHRAKGRGP 223

RESULT 10
 AAU35693
 ID ID AAU35693 standard; protein; 238 AA.
 XX
 AC AAU35693;
 XX
 XX 14-FEB-2002 (first entry)
 XX Helicobacter pylori cellular proliferation protein #6.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 XX Helicobacter pylori.
 OS
 OS Helicobacter pylori.
 XX
 PN WO200170955-A2.
 XX
 PN 27-SEP-2001.
 PD
 XX
 XX 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 XX WPI; 2001-611495/70.
 DR N-PSDB; AAS53552.
 DR
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 XX Example 3; SEQ ID NO 11286; 51lpp; English.
 PS
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 238 AA;
 SQ

Qy	1	VKLTPKEQKELLYYAGEVARRRKAAGLKLNOPEAIYISAHIMDEARRGKKTVLAQMEE	60
Dd	1	MKLTPKELDKLMLHYAGELAKRKEKGIKLVNVEAVALLSAHIMEEARAGKKTAAELMQE	60
Qy	61	CMHFLLKDEVMVPGVGNMVPDLGVZATPDGFKLVTWNWPIEPDEHFPAKEVVKFGCKDIE	120
Dd	61	GRTLLKPDDVDMDGVASMIHEVGIEIAMPDPDGFKLVTVHTPIEANGKLVPFGEI-FLKNEDIT	119
Qy	121	LNACKEVTELEVTNEGPKSLHVGSHPFHFEANKALKPDREKAYGKRLLDIPSGNTULRIAG	180
Dd	120	INEGKAVSVKVNVGDRPVQIGSHFFHFEVNRCLDFDREKTFGRLLDIASGTAVRFEPG	179
Qy	181	QTRKVLQILPLGGSKKVIWMGLVNNIADERHKHKALKDAKSHGF	224
Dd	180	EESVELIDIGNNRRIFGFNALVDROQDNESKKIALHRAKERGF	223
RESULT	11		
AAM49644	ID	AAM49644 standard; protein; 238 AA.	
XX	AC	AAM49644;	
XX	DT	17-MAY-2002 (first entry)	
XX	DE	Urease A protein PA4.	
XX	KW	Urease A; vaccine; T cell-stimulating peptide; antibacterial; antiulcer;	
XX	KW	MHC; major histocompatibility complex; epitope; antiinflammatory;	
XX	KW	Cytotoxicity; virucide; protozoicide; immunosuppressive; pathogen; tumour;	
XX	OS	autoimmune disease; Helicobacter pylori infection.	
XX	OS	Unidentified.	
XX	FN	WO200202141-A2.	
XX	FD	10-JAN-2002.	
XX	PF	04-JUL-2001; 2001WO-EP007656.	
XX	PR	05-JUL-2000; 2000DE-01032538.	
XX	FA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.	
XX	PI	Aebischer A, Bumann D, Lucas B, Meyer TF;	
XX	DR	WPI; 2002-164497/21.	
XX	PT	Preparing vaccine containing T cell-stimulating peptides, useful particularly for protection against Helicobacter pylori, by selecting epitopes from target proteins.	
XX	PS	Example 1; Fig 3; 36pp; German.	
XX	CC	This invention describes a novel method for the preparation of a vaccine based on T cell-stimulating peptide sequences. The method comprises (1) identifying potential MHC (major histocompatibility complex)-binding and T cell-reactive epitopes in a target protein; (2) eliminating cross-reactive epitopes, and (3) verifying remaining epitopes by determining their activity in vitro and/or in vivo. The products of the invention have antibacterial, antiulcer, antiinflammatory, cytostatic, virucide, protozoicide and immunosuppressive activity. The vaccine is especially used to protect against infection by Helicobacter pylori, but more generally can be prepared from proteins from any pathogen (virus, bacterium, protozoan) or tumour, or associated with autoimmune disease. Synthetic epitopes that are more immunogenic than natural sequences can be identified, and the vaccine may have increased immunogenicity (when formulated with adjuvant) or immunotolerance (formulated without an adjuvant). This sequence represents the urease A protein PA4 described in the invention	
SQ		Sequence 238 AA;	

Query Match	52.2%	Score 616;	DB 4;	Length 238;
Best Local Similarity	52.7%	Pred. No. 2.2e-19;		

Query Match 52.2%; Score 616; DB 5; Length 238;
Best Local Similarity 52.7%; Pred. No. 2.2e-19;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEOKFLLYAGVAVARKKAEGLKNQPEAIAVISAHIMDEARRGKKTVAOLMEE 60
DB 1 MKLTPKELDKLMLHYAGELAKRKEGKILNVVEAVALISAHIMEARAGKKTAAELMQE 60
QY 61 CMHFLKDEVMGPNVMDLGVATPDGKTLVTVMNPIEDPHFKAGEVKGCDKDIE 120
DB 61 GRTLLKPDVMDGVASMIHEVGIEAMPDGTGLVTVTHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALFDPREKAYGKRLDIPSGNTLRIGAG 180
DB 120 INEGKAVSVKVNVDGPRVQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179
QY 181 QTRKVLQIPLGSKKVIKMGVGLVNNIADERHKKALDKAKSHGF 224
DB 180 EKSVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 223
RESULT 12
AEB55115
ID AEB55115 standard; protein; 245 AA.
XX
AC AEB55115;
DT 22-SEP-2005 (first entry)
XX
DE Hylicobacter pylori Urase A protein fragment SEQ ID 4.
XX
KW pHUR3; vaccine; helicobacter pylori infection; urease;
KW immune stimulation; antibacterial; immunostimulant; antiulcer;
KW cytostatic; antiinflammatory; gastrointestinal-gen.; gastritis;
KW peptic ulcer; carcinoma.
XX
OS Helicobacter pylori.
OS Salmonella typhimurium.
XX
XX
PN W0200132014-A2.
XX
PD 10-MAY-2001.
XX
XX 01-NOV-2000; 2000WO-US030191.
XX
XX 01-NOV-1999; 99US-00431705.
XX
XX (ORAV-) ORAVAX INC.
XX
XX Kleanthous H, Londono-Arcila P, Freeman D;
XX
XX WPI; 2001-343379/36.
DR N-PSDB; AEB55112, AEB55130.
XX
XX Inducing an immune response against Helicobacter in mammals, useful for
PT treating Helicobacter induced gastroduodenal diseases.
XX
XX Disclosure; SEQ ID NO 4; 63pp; English.
XX
XX The invention relates to inducing an immune response against Helicobacter
CC in a mammal, comprising mucosally administering to the mammal an
CC attenuated Salmonella vector containing a nucleic acid molecule encoding
CC a Helicobacter antigen, and parenterally administering to the mammal a
CC Helicobacter antigen. Also included is an attenuated Salmonella vector
CC comprising a nucleic acid molecule encoding a Helicobacter antigen. The
CC Helicobacter antigen is a urease, a urease subunit, or its immunogenic
CC fragment (encoded by the ureA and ureB genes). The mammal is at risk of
CC having but does not have Helicobacter infection or has a Helicobacter
CC infection. The attenuated Salmonella vector further comprises an htrA or
CC ntrB promoter. The vector can be used in inducing an immune response
CC against Helicobacter in a mammal. The vector can be used to treat
CC Helicobacter infection. The vector and the method can be used to treat

CC Helicobacter induced gastroduodenal diseases, including acute, chronic or
CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
CC ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was
CC constructed expressing UreA and UreB and including an E. coli AmpR gene.
CC The present sequence represents a protein or peptide encoded by the pHUR3
CC plasmid. NOTE: pHUR3 is represented by both AEB55112 and AEB55130, the
CC peptides expressed by pHUR3 are shown in the sequence listing to be split
CC between AEB55112 and AEB55130, yet Figure 4 shows all the peptides being
CC expressed by the one sequence, therefore all encoded peptides are cross-
XX referenced to both sequences.
SQ Sequence 245 AA;
Query Match 52.2%; Score 616; DB 4; Length 245;
Best Local Similarity 52.7%; Pred. No. 2.4e-19;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEOKFLLYAGVAVARKKAEGLKNQPEAIAVISAHIMDEARRGKKTVAOLMEE 60
DB 8 MKLTPKELDKLMLHYAGELAKRKEGKILNVVEAVALISAHIMEARAGKKTAAELMQE 67
QY 61 CMHFLKDEVMGPNVMDLGVATPDGKTLVTVMNPIEDPHFKAGEVKGCDKDIE 120
DB 61 GRTLLKPDVMDGVASMIHEVGIEAMPDGTGLVTVTHTPIEANGKLVPGEL-FLKNEDIT 126
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALFDPREKAYGKRLDIPSGNTLRIGAG 180
DB 127 INEGKAVSVKVNVDGPRVQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 186
QY 181 QTRKVLQIPLGSKKVIKMGVGLVNNIADERHKKALDKAKSHGF 224
DB 187 EKSVELIDIGNRRIFGFNALVDROADNESKIALHRAKERGF 230
RESULT 13
ADS09178
ID ADS09178 standard; protein; 238 AA.
XX
AC ADS09178;
XX
DT 16-DEC-2004 (first entry)
XX
DE H. pylori urease alpha beta subunit.
XX
KW Urease; alpha subunit; UreB; gastric ulcer; peptic ulcer; gastritis;
KW stomach cancer; vaccine; antibody; immune reaction.
XX
OS Helicobacter pylori.
XX
XX JP2004261080-A.
XX
PD 24-SEP-2004.
XX
XX 28-FEB-2003; 2003JP-00054654.
XX
XX 28-FEB-2003; 2003JP-00054654.
XX
XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
XX (FUKU-) FUKUYAMA RINSHO KENSA CENT KK.
XX
XX WPI; 2004-665475/65.
XX
XX Protein having specific region of beta-subunit urease of Helicobacter
PT pylori, which generates antibodies that suppress proliferation and urease
PT activity of bacteria, useful for preventing Helicobacter pylori
PT infection.
XX
XX Disclosure; SEQ ID NO 3; 26pp; Japanese.
XX
XX The invention relates to a protein having a specific region of primary
CC structure of the beta-subunit Helicobacter pylori urease, appearing as
CC ADS09176 (amino acids 201-338 of the beta subunit) or a sequence which
CC generates antibodies that suppresses proliferation of H. pylori and which

CC suppresses urease activity of *H. pylori* by an immune reaction. Also
 CC included are a gene which encodes the urease fragment, a vaccine against
 CC *H. pylori* comprising the urease fragment, a transformed cell transformed
 CC by introducing the gene, an antibody which recognises the urease fragment
 CC and transforming a plant by introducing the gene into the plant and
 CC making it express. The urease fragment is useful for producing blood
 CC serum containing antibodies which suppresses proliferation of *H. pylori*
 CC and which suppresses urease activity of *H. pylori* which involves
 CC immunising a living organism (e.g. a human, cow or hen) with the urease
 CC fragment. The fragment may be purified from cow's milk or hen eggs from
 CC transgenic animals expressing the fragment. *H. pylori* is thought to be a
 CC cause of gastric and peptic ulcers, gastritis and may be implicated in
 CC stomach cancer. The urease enzyme is thought to be responsible for
 CC maintaining the bacterium in the acid conditions of the stomach. The
 CC present sequence is the *H. pylori* urease alpha subunit.

SQ Sequence 238 AA;
 Query Match 51.9%; Score 613; DB 8; Length 238;
 Best Local Similarity 52.2%; Pred. No. 3e-19;
 Matches 117; Conservative 44; Mismatches 62; Indels 1; Gaps 1;
 Qy 1 VKLTPKEQKELLYAGVAVKRAEGLKLNQPEAIAYISAHIMDEARRGKKTVAOLMER 60
 Db 1 MKLTPKELDKMLHYAGELAKKREKGIKLVYEAVALISAHIMEEARAGKKTAAELMQE 60
 Qy 61 CMHFLKDEVMGPNVGMVDPDLGVEATPDGTLKLVTVNWPIDPDEHFKAGEVKFGCDKDIE 120
 Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INEGKAVSVKVPVGDPRVQIGSHFHFPEVNRCLDFDREKTFGKRLDIAAGTAVRPFPG 179
 Qy 181 QTRKVLPIPGSKKVGNGLVNNIADERHKHAKLDKAKSHGF 224
 Db 180 EKSVELIDIGGNRRIFGNALVDROADNESKKTALHRAKERGF 223

RESULT 14

ADM28643
 ID ADM28643 standard; protein; 238 AA.
 AC ADM28643;
 XX
 XX 20-MAY-2004 (first entry)
 DT
 DE
 KW Helicobacter pylori urease alpha subunit protein SeqID 2.
 KW immunogenic; urease; vaccine; passive immunisation; diagnostic;
 KW antibacterial; antibody.
 XX
 OS Helicobacter pylori.
 XX
 XX JP2004041084-A.
 XX
 PD 12-FEB-2004.
 XX
 XX 11-JUL-2002; 2002JP-00203221.
 PF
 XX 11-JUL-2002; 2002JP-00203221.
 PR
 XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 XX
 XX WPI; 2004-162021/16.
 DR
 XX New peptide useful as a vaccine against Helicobacter pylori induces an
 PT antibody response against the bacterial urease.
 XX

PS Disclosure; SEQ ID NO 2; 19pp; Japanese.
 XX
 XX This invention relates to a novel immunogenic peptide that induces
 CC production of antibodies against the Helicobacter pylori (*H. pylori*)

CC urease protein. Specifically, it refers to a peptide antigen capable of
 CC inducing an antibody response, such that it can be used to develop a
 CC vaccine against *H. pylori*. The present invention describes generating
 CC monoclonal antibodies against the urease using hybridoma techniques that
 CC are useful for passive immunisation and diagnostic techniques.
 CC Furthermore, the antibacterial peptide can be used to immunise an animal
 CC in order to produce antibodies that are present within a functional food
 CC such as milk or eggs. This polypeptide sequence is the *H. pylori* urease
 CC alpha subunit protein of the invention.

SQ Sequence 238 AA;

Query Match 51.6%; Score 609; DB 8; Length 238;
 Best Local Similarity 52.2%; Pred. No. 4.7e-19;
 Matches 117; Conservative 43; Mismatches 63; Indels 1; Gaps 1;
 Qy 1 VKLTPKEQKELLYAGVAVKRAEGLKLNQPEAIAYISAHIMDEARRGKKTVAOLMER 60
 Db 1 MKLTPKELDKMLHYAGELAKKREKGIKLVYEAVALISAHIMEEARAGKKTAAELMQE 60
 Qy 61 CMHFLKDEVMGPNVGMVDPDLGVEATPDGTLKLVTVNWPIDPDEHFKAGEVKFGCDKDIE 120
 Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INEGKAVSVKVPVGDPRVQIGSHFHFPEVNRCLDFDREKTFGKRLDIAAGTAVRPFPG 179
 Qy 181 QTRKVLPIPGSKKVGNGLVNNIADERHKHAKLDKAKSHGF 224
 Db 180 EKSVELIDIGGNRRIFGNALVDROADNESKKTALHRAKERGF 223

RESULT 15

AAR67375
 ID AAR67375 standard; protein; 238 AA.
 XX
 AC AAR67375;
 XX
 XX 25-MAR-2003 (revised)
 DT 22-JUN-1995 (first entry)
 XX
 XX *H. pylori* ureA urease.
 DE
 XX Urease; ureA gene; immunogen; vaccine; diagnostic; Helicobacter felis.
 KW
 XX Proteus mirabilis.
 OS
 XX WO9426901-A1.
 XX
 PD 24-NOV-1994.
 XX
 XX 19-MAY-1994; 94WO-EP001625.
 PF
 XX 19-MAY-1993; 93EP-00401309.
 PR 19-NOV-1993; 93WO-EP003259.
 XX
 XX (INSP) INST PASTEUR.
 PA (INRM) INST NAT SANTE & RECH MEDICALE.
 XX
 XX Labigne A, Suerbaum S, Ferrero R, Thiberge J;
 XX WPI; 1995-006797/01.
 DR
 XX DNA from Helicobacter pylori and Helicobacter felis - used to develop
 PT prods. for detection, treatment and prevention of Helicobacter infection.
 XX
 XX Disclosure; Fig 4i-iii; 168pp; English.
 XX

CC The sequence of the Helicobacter felis urease ureA gene product (given in
 CC AAR67371) was compared to ureases of Helicobacter pylori (AAR67375),
 CC Proteus mirabilis (AAR67376) and jack bean urease (AAR67377) and regions
 CC of homology were identified. (Updated on 25-MAR-2003 to correct PN

```

CC field.)
XX
SQ Sequence 238 AA;
    Query Match      51.5%; Score 608; DB 2; Length 238;
    Best Local Similarity 52.2%; Pred. No. 5, 2e-19;
    Matches 117; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

Qy 1 VKLTPKQEKFLYYAGVARRKKAEGKLNQPEAIAIYISAHIMDEARRGKKTVAQLMEE 60
    :||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKLTPKELDKLMLHYAGELAKKKEKGIKLNVEAVALISAHIMEARAGKKTAAELMQE 60
    :||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 61 CMHFLKKDEVMPGCVGMVDPDLGVEATPDDGTCLVTVNWPIDPDEHFKAQEVKFGCDKDIE 120
    :||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 GRTLLKPDVNDMGVASMIEHVGIEAMFPDGTCLVTVHTPIEANGKLVPEL-FLKNEDIT 119
    :||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
    :||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 INEGKAVSVKKNVGDPRVQIGSHFHFVFNRCLDPDRKTFHKKRLDIASGTAVRFPFG 179
    :||||| : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 181 QTRKVQLIPLGSKKVIQMGNLVNNIADERHKKHKKALDKAKSHGF 224
    : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 EEKSVELIDIGNRRIFGFNALVDRQADNESKKTALHRAKERGF 223
    : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: August 9, 2006, 21:23:50
 Job time : 198 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:29:12 ; Search time 50 Seconds
(without alignments)
395.638 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADERHKHAKDCAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA.*
1: /EMC Celerra_IDS3/prodata/2/iaa/5 COMB.pep.*
2: /EMC Celerra_IDS3/prodata/2/iaa/6 COMB.pep.*
3: /EMC Celerra_IDS3/prodata/2/iaa/7 COMB.pep.*
4: /EMC Celerra_IDS3/prodata/2/iaa/H COMB.pep.*
5: /EMC Celerra_IDS3/prodata/2/iaa/PCTUS COMB.pep.*
6: /EMC Celerra_IDS3/prodata/2/iaa/RE COMB.pep.*
7: /EMC Celerra_IDS3/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638	54.0	234	2	US-09-338-920B-11
2	616	52.2	238	1	Sequence 11, Appli
3	616	52.2	238	1	Sequence 2, Appli
4	616	52.2	238	5	Sequence 5, Appli
5	616	52.2	238	5	Sequence 2, Appli
6	611	51.7	237	1	Sequence 4, Appli
7	611	51.7	237	1	Sequence 22, Appli
8	611	51.7	237	2	Sequence 22, Appli
9	609	51.6	238	2	Sequence 7, Appli
10	579	49.0	228	2	Sequence 154, App
11	578	48.9	237	1	Sequence 20, Appli
12	578	48.9	237	2	Sequence 20, Appli
13	578	48.9	237	2	Sequence 20, Appli
14	572	48.4	237	2	Sequence 9, Appli
15	554	46.9	213	2	Sequence 13, Appli
16	451	38.2	840	1	Sequence 25, Appli
17	451	38.2	840	2	Sequence 25, Appli
18	451	38.2	840	2	Sequence 25, Appli
19	287	24.3	100	1	Sequence 3, Appli
20	287	24.3	100	1	Sequence 3, Appli
21	287	24.3	103	2	Sequence 3, Appli
22	275	23.3	99	1	Sequence 9261, Ap
23	273	23.1	145	2	Sequence 1, Appli
24	272	23.0	107	2	Sequence 4982, Ap
25	270	22.9	132	2	Sequence 6101, Ap
26	268	22.7	103	2	Sequence 5049, Ap

27 267 22.6 100 1 US-08-467-822-23 Sequence 23, Appli
28 267 22.6 100 2 US-08-432-697-23 Sequence 23, Appli
29 267 22.6 100 2 US-08-466-248-23 Sequence 23, Appli
30 259 21.9 125 2 US-09-252-991A-26884 Sequence 26884, A
31 256 21.7 100 2 US-09-602-777A-18 Sequence 18, Appli
32 250 21.2 111 2 US-09-710-279-116 Sequence 116, App
33 243 20.6 153 2 US-09-543-681A-6085 Sequence 6085, Ap
34 234 19.8 106 1 US-07-732-242C-2 Sequence 2, Appli
35 229 19.4 119 2 US-09-328-352A-5915 Sequence 5915, Ap
36 221 18.7 109 1 US-08-467-822-24 Sequence 24, Appli
37 221 18.7 109 2 US-08-432-697-24 Sequence 24, Appli
38 221 18.7 109 2 US-08-466-248-24 Sequence 24, Appli
39 219 18.5 106 1 US-08-967-513-4 Sequence 4, Appli
40 219 18.5 106 1 US-08-687-645B-4 Sequence 4, Appli
41 218 18.5 109 2 US-09-489-039A-9266 Sequence 9266, Ap
42 217 18.4 162 2 US-09-602-777A-10 Sequence 10, Appli
43 197 16.7 137 2 US-09-252-991A-26886 Sequence 26886, A
44 171 14.5 289 2 US-09-134-001C-4415 Sequence 4415, Ap
45 167 14.1 791 2 US-10-104-047-2307 Sequence 2307, Ap

ALIGNMENTS

RESULT 1
US-09-338-920B-11
; Sequence 11, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-09-338-920B-11

Query Match 54.0%; Score 638; DB 2; Length 234;
Best Local Similarity 54.8%; Pred. No. 6.9e-23;
Matches 126; Conservative 41; Mismatches 55; Indels 8; Gaps 5;
Qy 1 VKLTPKEQKFLLYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAOLMEE 60
Db 1 MKLTPKELDKMLHAYAGELAKQKAKGKLNTEAVALISAHVMEARAGKKSVDLMQE 60
Qy 61 CMHFLKCKDEVNPGVGNMVPDLGVEATFPDGTGLVTNNWPIEP-DEHFKAGEV--KFGCD 117
Db 61 GRTLLKADDVMPGVAHMLHEVGIEAGFPDGTGLVTIHTPEAGSDKLAPEVILK---NE 117
Qy 118 DIENAGKEVTELEVNTNPGKSLHVGSHFHPFEAKALKFDPREKAYGKGLDIPSGNTLRI 177
Db 118 DITLNAGKHAQLKVKNNKGRPVQVQSHFHPFEVNNKLLDFDREKAYGKGLDIASGTAVER 177
Qy 178 GAGQTRKQVLIPLGSKKVIWNGLVNNIADERHKHAKDLDKAK-SH-GFI 225
Db 178 EPGBEKTVELDIGNKKRIYGFNALVDROADHGDGKKLAKRAKKEHFGTI 227

RESULT 2
US-08-920-095-2
; Sequence 2, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.

```
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-920-095-2

Query Match      52.2%; Score 616; DB 1; Length 238;
Best Local Similarity 52.7%; Pred. No. 8.6e-22;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

Qy      1 VKLTPKEQKFLYYAGVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db      1 MKLTPKELDKLMLHYAGELAKRKEGKIKLYVEAVALLSAHIMEARAGKKTAAELMQE 60
Qy      61 CMHFLKDEVMVPGVGNMVPDLGVATFPDGTGLVTNNPIEPDEHFHFKAGEVKFGCDKIE 120
Db      61 GRTLKPDVMDGVASMIHEVGIEAMFPDGTGLVTVHTPIEANGKLVPEGL-FLKNEDIT 119
Qy      121 LNAGKEVTELEVTNEGPKSLHVGSHFHPPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
Db      120 INEGKAVSVKVNQVDRPVQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179
Qy      181 QTRKVQLIPLGGSKKVGIMGNLVNNIADERHKHKKALDKAKSHGF 224
Db      180 EKSVELIDIGNRRIFGNALVDRQADNESKIALHRAKERGF 223

; RESULT 3
; Sequence 5, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 5
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
; US-09-338-920B-5

Query Match      52.2%; Score 616; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 8.6e-22;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

Qy      1 VKLTPKEQKFLYYAGVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db      1 MKLTPKELDKLMLHYAGELAKRKEGKIKLYVEAVALLSAHIMEARAGKKTAAELMQE 60
Qy      61 CMHFLKDEVMVPGVGNMVPDLGVATFPDGTGLVTNNPIEPDEHFHFKAGEVKFGCDKIE 120
Db      61 GRTLKPDVMDGVASMIHEVGIEAMFPDGTGLVTVHTPIEANGKLVPEGL-FLKNEDIT 119
Qy      121 LNAGKEVTELEVTNEGPKSLHVGSHFHPPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
Db      120 INEGKAVSVKVNQVDRPVQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179
Qy      181 QTRKVQLIPLGGSKKVGIMGNLVNNIADERHKHKKALDKAKSHGF 224
Db      180 EKSVELIDIGNRRIFGNALVDRQADNESKIALHRAKERGF 223

; RESULT 4
; PCT-US96-05800-2
; Sequence 2, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: Oravax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-05800-2
```

APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

APPLICATION NUMBER: US 08/467,822
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/447,177
 FILING DATE: 19-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/432,697
 FILING DATE: 02-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03495.0137-02000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4400
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 237 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-467-822-22

```

; MOLECULE TYPE: protein
US-08-467-822-22

Query Match      51.7%; Score 611; DB 1; Length 237;
Best Local Similarity 52.7%; Pred. No. 1.5e-21;
Matches 118; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

Qy   1 VKLTPKSEQEKLPLYYAGEVAKRKKAAGLKLNQPEAIAIYSAHIMDARCKKTVAOLMEE 60
Db   1 MKLTPKELDK-LMHYAGELAKRKEGKIKLVNVEAVALIYSAHIMEBARAGKTKTAABLMOE 59

Qy   61 CMHFLLKDEVMYPGVGNVPDLGVEATPDDGTKLVTNNMPIEPDEHFKAAGSVKFGCDKDIE 120
Db   60 GRILLKPDDVMDGVASMIHEVGISAMPDGGTKLVTVHTPIEANGKVLVPGBEL-FLKNEDIT 118

Qy   121 LNAGKEVTELEVTNEGPKSLHVSGSHFFHFEANKALKFDREKAYGKRLEDIPSGNTLRIGAG 180
Db   119 INEGKKAVSVKVNVDGPVQVIGSHFFHFEVNRCLDPFREKTFGKELDIASGTAVERFPG 178

Qy   181 QTRKVQLPIPGGSKKVITGMGLNVNNIADERHHKHKALDKAKSHGF 224
Db   179 EEKSVELIDIGENRRIFGFNALVDROADNESKKIALHRAKERGF 222

```

Db 179 EEKSVELDIGENRIFGFNALVDRQADNESKIALHRAKGRGF 222

RESULT 7
US-08-432-697-22
; Sequence 22, Application US/08432697
; Patent No. 6248330

Db 1 MKLTPKELDKMLHYAGRLAERLARGVKLNTEAVALISGRVMEKARDGNKSVADLMQ 60
Qy 61 CMHFLKDEVMGPNMVPDLGVEATFPDGTGLVTNNPFIPEDEHFKAGEVKFGCDKDI 120
Db 61 GRTWLKXENVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT 119
Qy 121 LNAKEVTEVETNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIAG 180
Db 120 INAGKEAISLKVKNKGRPVQVQSHFHFPEVFNKLLDFDRAKSFCKRLDIASGTAVRFP 179
Qy 181 QTRKVLPIPLGSKKVGIMGNLVNNIADERHKKHAKDKAKSHGF 224
Db 180 EKSVELIDIGNKKRIYFNSLVDRQADAGKUGLKRKAKGKF 223

RESULT 12
US-08-432-697-20
; Sequence 20, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1-237
; OTHER INFORMATION: /note= "URE A - FIGURE 3."
US-08-432-697-20

Query Match 48.9%; Score 578; DB 2; Length 237;
Best Local Similarity 50.4%; Pred. No. 6.2e-20;
Matches 113; Conservative 43; Mismatches 67; Indels 1; Gaps 1;
Qy 1 VKLTPKEQKFLYYAGVARRKRAEGLKNOPEAIAYISAHIMDEARRGKKTVAQLME 60
Db 1 MKLTPKELDKMLHYAGRLAERLARGVKLNTEAVALISGRVMEKARDGNKSVADLMQ 60

Qy 61 CMHFLKDEVMGPNMVPDLGVEATFPDGTGLVTNNPFIPEDEHFKAGEVKFGCDKDI 120
Db 61 GRTWLKXENVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT 119
Qy 121 LNAKEVTEVETNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIAG 180
Db 120 INAGKEAISLKVKNKGRPVQVQSHFHFPEVFNKLLDFDRAKSFCKRLDIASGTAVRFP 179
Qy 181 QTRKVLPIPLGSKKVGIMGNLVNNIADERHKKHAKDKAKSHGF 224
Db 180 EKSVELIDIGNKKRIYFNSLVDRQADAGKUGLKRKAKGKF 223

RESULT 13
US-08-466-248-20
; Sequence 20, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1-237
; OTHER INFORMATION: /note= "URE A - FIGURE 3."
US-08-466-248-20

```
Query Match      48.9%; Score 578; DB 2; Length 237;
Best Local Similarity 50.4%; Pred. No. 6.2e-20;
Matches 113; Conservative 43; Mismatches 67; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKELLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKLTPKELDKMLHYAGRLAEERLARGVKLNYTEAVALISGRVMEKARDGNKSVADLMQE 60
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 61 CMHFLKXDEVMPGVGNMVPDLGVEATPDGKLTVTNNPIEPDEHFHFKAGEVKGCDKDIE 120
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 GRTWLKKNVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT 119
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDKREKAYGKRLDIPSGNTLRIGAG 180
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 INAGKEAISLVKKNKGRDPVQVGHFHFPEVKNKLLDFDRAKSFCKRLDIASGTAVRFPFG 179
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 181 QTRKVQLIPLGGSKVKVIGMGNLVNNIADERHKHKALDKAKSHGF 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 EKSVELIDIGGNKRIYGFNSLVDRQADADGKGLGLKRAKEKGF 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-338-920B-9
; Sequence 9, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Helicobacter felis
; US-09-338-920B-9
```

```
Query Match      48.4%; Score 572; DB 2; Length 237;
Best Local Similarity 50.0%; Pred. No. 1.2e-19;
Matches 112; Conservative 43; Mismatches 68; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKELLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKLTPKELDKMLHYAGRLAEERLARGVKLNYTEAVALISGRVMEKARDGNKSVADLMQE 60
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 61 CMHFLKXDEVMPGVGNMVPDLGVEATPDGKLTVTNNPIEPDEHFHFKAGEVKGCDKDIE 120
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 GRTWLKKNVMDGVASMIHEVGIEANFPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT 119
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDKREKAYGKRLDIPSGNTLRIGAG 180
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 INAGKEAISLVKKNKGRDPVQVGHFHFPEVKNKLLDFDRAKSFCKRLDIASGTAVRFPFG 179
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 181 QTRKVQLIPLGGSKVKVIGMGNLVNNIADERHKHKALDKAKSHGF 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 EKSVELIDIGGNKRIYGFNSLVDRQADADGKGLGLKRAKEKGF 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 15
US-09-338-920B-13
; Sequence 13, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
```

```
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Helicobacter mustelae
; US-09-338-920B-13

Query Match      46.9%; Score 554; DB 2; Length 213;
Best Local Similarity 52.3%; Pred. No. 7.1e-19;
Matches 113; Conservative 37; Mismatches 61; Indels 5; Gaps 2;

Qy 13 LYVAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMECHMFLKXDEVMP 72
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 LHYAGELAKKRXERGKLVNYVEAVALISMEIHAREGKKTVAQLMECHMFLKXDEVMP 60
   :|||||: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 73 GVGNNMVPDLGVEATPDGKLTVTNNPIEPDEHFHFKAGE--VKFGCDKDIELNAGKEVTEL 130
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 GVAALVHEVQIEVCFPDGTLTLLVTVHNPINNGKLFHGEFILK---DEDIVLNAGKEAIEV 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 131 EVTNEGPKSLHVGSHFHFPEANKALKPDKREKAYGKRLDIPSGNTLRIGAGTRKVQLIPL 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 118 KVSNGKDRPIQVGHFHFPEANKALKPDKREKAYGKRLDIPSGNTLRIGAGTRKVQLIPL 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 191 GGSKKVIGMGNLVNNIADERHKHKALDKAKSHGF 226
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 GGNQRIFGFNDLNGQVNEEDNKRKALAAKAKGFIK 213
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: August 9, 2006, 21:30:35
Job time : 51 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:29:51 ; Search time 184 Seconds
(without alignments)
568.949 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKQEKFLYYAGEVA.....ADERHKHALDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main.*

1: /EMC_Celerra_SID3S/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /EMC_Celerra_SID3S/prodata/2/pubpaa/US08_PUBCOMB.pep.*
3: /EMC_Celerra_SID3S/prodata/2/pubpaa/US09_PUBCOMB.pep.*
4: /EMC_Celerra_SID3S/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
5: /EMC_Celerra_SID3S/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
6: /EMC_Celerra_SID3S/prodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1181	100.0	226	3	US-09-904-994B-2
2	1178	99.7	226	3	US-09-904-994B-14
3	1174	99.4	226	3	US-09-904-994B-5
4	1174	99.4	226	3	US-09-904-994B-11
5	1170	99.1	226	3	US-09-904-994B-8
6	665	56.3	234	4	US-10-639-273-3
7	638	54.0	234	4	US-10-639-273-34
8	619	52.4	238	3	US-09-815-242-11468
9	616	52.4	238	3	US-10-335-977-8629
10	605	51.2	237	4	US-09-815-242-11286
11	582	49.3	228	4	US-10-639-273-36
12	579	49.0	228	4	US-10-335-977-8628
13	572	48.4	237	4	US-10-012-819-154
14	556	47.1	225	4	US-10-639-273-35
15	530	44.9	811	4	US-10-639-273-37
16	516	43.7	207	4	US-10-282-122A-70574
17	509	43.1	226	3	US-10-282-122A-70447
18	505	42.8	779	4	US-09-815-242-5448
19	493	41.7	207	4	US-10-282-122A-69134
20	492	41.7	207	4	US-10-282-122A-45038
21	490	41.5	261	4	US-10-282-122A-48209
22	487	41.2	209	4	US-10-282-122A-55507
23	467	39.5	228	4	US-10-282-122A-68251
24	462	39.1	840	4	US-10-282-122A-50077
25	462	39.1	840	5	US-10-621-833-7
26	462	39.1	840	5	US-10-731-877-1
27	462	39.1	840	6	US-11-046-271-1

28	436	36.9	843	4	US-10-437-963-124057	Sequence 124057,
29	430	36.4	171	4	US-10-335-977-8627	Sequence 8627, Ap
30	415	35.1	837	4	US-10-424-599-254635	Sequence 254635, A
31	408	34.5	227	4	US-10-156-761-10253	Sequence 10253, A
32	382	32.3	837	4	US-10-418-962-2	Sequence 2, Appli
33	317	26.8	227	4	US-10-425-115-201340	Sequence 201340,
34	291	24.6	121	4	US-10-282-122A-49772	Sequence 49772, A
35	287	24.3	100	3	US-09-815-242-11745	Sequence 11745, A
36	287	24.3	100	4	US-10-156-761-14633	Sequence 14633, A
37	287	24.3	100	4	US-10-282-122A-60250	Sequence 60250, A
38	284	24.0	100	4	US-10-282-122A-62539	Sequence 62539, A
39	284	24.0	100	4	US-10-282-122A-64616	Sequence 64616, A
40	283	24.0	100	4	US-10-282-122A-56822	Sequence 56822, A
41	283	24.0	100	4	US-10-282-122A-56846	Sequence 56846, A
42	276	23.4	100	4	US-10-282-122A-69398	Sequence 69398, A
43	273	23.1	100	3	US-09-815-242-11047	Sequence 11047, A
44	273	23.1	100	4	US-10-282-122A-58211	Sequence 58211, A
45	273	23.1	145	4	US-10-724-972A-6468	Sequence 6468, Ap

ALIGNMENTS

RESULT 1

US-09-904-994B-2
; Sequence 2, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-2

Query Match		100.0%;	Score 1181;	DB 3;	Length 226;
Best Local Similarity		100.0%;	Pred. No. 3.3e-42;		
Matches 226;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VKLTPKQEKFLYYAGVARKKAEGLKNOPEAIYISAHIMDEARRGKKTVAOLMEE	60		
Db	1	VKLTPKQEKFLYYAGVARKKAEGLKNOPEAIYISAHIMDEARRGKKTVAOLMEE	60		
Qy	61	CMHFLKQDEVMVPGVGNMVPDLGVEATFPDGTKLVTNNWPIEPDEHFKAGEVKGCDKDIE	120		
Db	61	CMHFLKQDEVMVPGVGNMVPDLGVEATFPDGTKLVTNNWPIEPDEHFKAGEVKGCDKDIE	120		
Qy	121	LNAGKEVTELEVTNKGKSLHVGSHFHPFEANKALKFDEKAYGKRLDIPSGNTLIRIGAG	180		
Db	121	LNAGKEVTELEVTNKGKSLHVGSHFHPFEANKALKFDEKAYGKRLDIPSGNTLIRIGAG	180		
Qy	181	QTRKVLPLGSKKVI GNGLVNNIADERHKHALDKAKSHGFIK	226		
Db	181	QTRKVLPLGSKKVI GNGLVNNIADERHKHALDKAKSHGFIK	226		

RESULT 2

US-09-904-994B-14
; Sequence 14, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-14

Query Match          99.7%; Score 1178; DB 3; Length 226;
Best Local Similarity 99.6%; Pred. No. 4.4e-42;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Qy 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 3
US-09-904-994B-5
; Sequence 5, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-5

Query Match          99.4%; Score 1174; DB 3; Length 226;
Best Local Similarity 99.1%; Pred. No. 6.6e-42;
Matches 224; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Qy 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 4
US-09-904-994B-11
; Sequence 11, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
```

```
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-11

Query Match          99.4%; Score 1174; DB 3; Length 226;
Best Local Similarity 99.1%; Pred. No. 6.6e-42;
Matches 224; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Qy 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 5
US-09-904-994B-8
; Sequence 8, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-8

Query Match          99.1%; Score 1170; DB 3; Length 226;
Best Local Similarity 98.7%; Pred. No. 9.8e-42;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Qy 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226

RESULT 6
US-09-904-994B-8
; Sequence 8, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-8

Query Match          99.1%; Score 1170; DB 3; Length 226;
Best Local Similarity 98.7%; Pred. No. 9.8e-42;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGVARRKKAEGKLKNOPEAIAYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Db 61 CMHFLKKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNNWPIEPDEHFKAGEVKFGCDKDIE 120
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Qy 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
Db 181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKKALDKAKSHGFIK 226
```

US-10-639-273-3
; Sequence 3, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-3

Query Match 56.3%; Score 665; DB 4; Length 234;
Best Local Similarity 57.3%; Pred. No. 5.6e-20;
Matches 130; Conservative 39; Mismatches 52; Indels 6; Gaps 3;

Qy 1 VKLTPKEQKFLYYAGVAVKRAEGLKLNQPAIAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRANGKLVNTEAVALISAHVMEERAGKKSVDLMQE 60

Qy 61 CMHFLKKDEVMGVMVDPDLGVEATPPDGTGLVTVNWPPIEPDEHFKA-GEV--KFGCDK 117
Db 61 GRTLLKADDDVMPGVAMHIEVGIEAGFPDGTGLVTIHTPVEAGSDKLAPGEVILK--NE 117

Qy 118 DIELNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRI 177
Db 118 DITLNAGKHAQVLKVNKGDRPVQVGVSHFHFPEVFNKLLDFDRKAYGKRLDIASGTAVRF 177

Qy 178 GAGOTRQVQLIPLGGSKKVIKMGVNLVNIADERHKHAKLDKAK-SH-GFI 225
Db 178 EPGEKTVELIDIGGNKRIYGFNALVDRQADHDGKKLAKRAKEKHGFI 227

RESULT 7
US-10-639-273-34
; Sequence 34, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 34
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter helmannii
US-10-639-273-34

Query Match 54.0%; Score 638; DB 4; Length 234;
Best Local Similarity 54.8%; Pred. No. 8.1e-19;
Matches 126; Conservative 41; Mismatches 55; Indels 8; Gaps 5;

Qy 1 VKLTPKEQKFLYYAGVAVKRAEGLKLNQPAIAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKQKAGIKLVNTEAVALISAHVMEERAGKKSVDLMQE 60

Qy 61 CMHFLKKDEVMGVMVDPDLGVEATPPDGTGLVTVNWPPIEP-DEHFKA-GEV--KFGCDK 117
Db 61 GRTLLKADDDVMPGVAMHIEVGIEAGFPDGTGLVTIHTPVEAGSDKLAPGEVILK--NE 117

Qy 118 DIELNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRI 177
Db 118 DITLNAGKHAQVLKVNKGDRPVQVGVSHFHFPEVFNKLLDFDRKAYGKRLDIASGTAVRF 177

Qy 178 GAGOTRQVQLIPLGGSKKVIKMGVNLVNIADERHKHAKLDKAK-SH-GFI 225
Db 178 EPGEKTVELIDIGGNKRIYGFNALVDRQADHDGKKLAKRAKEKHGFI 227

RESULT 8
US-09-815-242-11468
; Sequence 11468, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11468
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11468

Query Match 52.4%; Score 619; DB 3; Length 238;
Best Local Similarity 53.1%; Pred. No. 5.6e-18;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGVAVKRAEGLKLNQPAIAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKEKGIKLVNTEAVALISAHIMDEARRGKKTAAELMQE 60

Qy 61 CMHFLKKDEVMGVMVDPDLGVEATPPDGTGLVTVNWPPIEPDEHFKA-GEV--KFGCDK 120
Db 61 GRTLLKADDDVMPGVAMHIEVGIEAGFPDGTGLVTIHTPVEAGSDKLAPGEVILK--NE 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRI 180
Db 120 INEGKAVSVKVNKGDRPVQVGVSHFHFPEVFNKLLDFDRKAYGKRLDIASGTAVRF 179

Qy 181 QTRVQLIPLGGSKKVIKMGVNLVNIADERHKHAKLDKAKSHGF 224
Db 180 BEKSVELIDIGGNRRIFGFNALVDRQADNESKKIALHRAKERGF 223

RESULT 9
 US-10-335-977-8629
 ; Sequence 8629, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 10031
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: Windows NT 4.0
 ; SOFTWARE: UNIX
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/335,977
 ; FILING DATE: 30-Dec-2002
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 08/993,002
 ; FILING DATE: 17-DEC-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mandragoras, Amy E.
 ; REGISTRATION NUMBER: 36,207
 ; REFERENCE/DOCKET NUMBER: GTN-018
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)742-4214
 ; INFORMATION FOR SEQ ID NO: 8629:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 238 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Helicobacter pylori
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...238
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 8629:
 US-10-335-977-8629
 Query Match 52.4%; Score 619; DB 4; Length 238;
 Best Local Similarity 53.1%; Pred. No. 5.6e-18;
 Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;
 QY 1 VKLTPKEQKFLYYAGVARKKAEGLKLNQPEALAYISAHIMDEARRGKKTVAOLMEE 60
 Db 1 MKLTPKELDKMLHYAGELAKRKEGKIKLNTVEAVALISAHIMEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTKLVTNNWPIEDPHFKAGEVFGCDKQIE 120
 Db 61 GRTLLKPDVDMGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELVNTEGPKSLHVGSHFHPFEANKALFDEKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INEGKAVSVKVNVDGPRVQIGSHGFHPFEVNRCLDFDREKTFGKRLDIASGTAVRFEFG 179
 QY 181 QTRKVLIPLGSKKVIKMGVGNVNIADERHKKHAKDIAKSHGF 224
 Db 180 EKSVELIDIGNNRIFGNALVDRQADNESKIALHRAKERGF 223

RESULT 10
 US-09-815-242-11286
 ; Sequence 11286, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11286
 ; LENGTH: 238
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 US-09-815-242-11286
 Query Match 52.2%; Score 616; DB 3; Length 238;
 Best Local Similarity 52.7%; Pred. No. 7.5e-18;
 Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
 QY 1 VKLTPKEQKFLYYAGVARKKAEGLKLNQPEALAYISAHIMDEARRGKKTVAOLMEE 60
 Db 1 MKLTPKELDKMLHYAGELAKRKEGKIKLNTVEAVALISAHIMEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTKLVTNNWPIEDPHFKAGEVFGCDKQIE 120
 Db 61 GRTLLKPDVDMGVASMIHEVGIEAMFPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELVNTEGPKSLHVGSHFHPFEANKALFDEKAYGKRLDIPSGNTLRIGAG 180
 Db 120 INEGKAVSVKVNVDGPRVQIGSHGFHPFEVNRCLDFDREKTFGKRLDIASGTAVRFEFG 179
 QY 181 QTRKVLIPLGSKKVIKMGVGNVNIADERHKKHAKDIAKSHGF 224
 Db 180 EKSVELIDIGNNRIFGNALVDRQADNESKIALHRAKERGF 223
 RESULT 11
 US-10-639-273-36
 ; Sequence 36, Application US/10639273
 ; Publication No. US20040142343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chang, Yung-Fu
 ; APPLICANT: Simpson, Kenneth W
 ; APPLICANT: Zhu, Jiaqian
 ; TITLE OF INVENTION: HELICOBACTER BIZZOZERRONI UREASE GENES AND THEIR USES
 ; FILE REFERENCE: 19603/3881
 ; CURRENT APPLICATION NUMBER: US/10/639,273

RESULT 14

US-10-639-273-35
; Sequence 35, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; PRIOR FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-10-639-273-35

Query Match 48.4%; Score 572; DB 4; Length 237;

Best Local Similarity 50.0%; Pred. No. 5.8e-16;

Matches 112; Conservative 43; Mismatches 68; Indels 1; Gaps 1;

Qy	1	VKLTPKEQKELLYAGVARRKKAEGKLKNOPEAIAVISAHIMDEARRGKKTVAQLMEE	60
Db	1	MKLTPKELDKLMLHYAGLAEALARGVKLNTYEAVALISGRVMEKARDGNKSVADLMQE	60
Qy	61	CMHFLKDEVMGPGVGNMVPDLGVEATFPDGTKLVTNNWPIEPDEHFKAEGVFCGCDKDIE	120
Db	61	GRTWLKKENVMGVSAMHEVGEIANFPDGTKLVTIHPVEDNGKLAFGEV-FLKNEDIT	119
Qy	121	LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYKGRLDIPSGNTLRLIG	180
Db	120	INAGKEAISLVKNKGRDPVQVGSHFHFPEVKNKLLDFDRAKSFCKRLDIASGTAVRFPFG	179
Qy	181	QTRKVQLIPLGGSKKVTGMNGLVNNIADERHKKALDKAKSHGF	224
Db	180	EKSVELIDIGGNKRIYGNLSLVRQADADGKGLGKRAKEKGF	223

RESULT 15

US-10-639-273-37
; Sequence 37, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Helicobacter hepaticus
US-10-639-273-37

Query Match 47.1%; Score 556; DB 4; Length 225;

Best Local Similarity 49.3%; Pred. No. 2.5e-15;

Matches 113; Conservative 45; Mismatches 62; Indels 9; Gaps 6;

Qy	1	VKLTPKEQKELLYAGVARRKKAEGKLKNOPEAIAVISAHIMDEARRGKKTVAQLMEE	60
----	---	-----------------------------------------------------------	----

Db	1	MRLTPKELDKLMLHYAGALAKSRKRGIKLNTVSTALISMEIMELAREGNKSVaelMQQ	60
Qy	61	CMHFLKDEVMGPGVGNMVPDLGVEATFPDGTKLVTNNWPIEPDEHFKAEG--VKFGCDKD	118
Db	61	GREILKADVMGEGVASMNEVQVSPDGTKLVTIHNPIEDNGKLTPEGYILK---DED	117
Qy	119	IELNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYKGRLDIPSGNTLRLIG	178
Db	118	ITLNANKESISIKVTHKGRPIQVGSHFHFPEVKNKLLDFDRAQAFGKRLDIASTSVRFE	177
Qy	179	AGOTRKVOLIPLGSKKVTGMNGLVNNIADERHKKAL-DKAKSHGFI	225
Db	178	PGEKNVNLIDFGGKOKIIGFNDLTNAHI-NKENKEQCLANAAQKH-FI	224

Search completed: August 9, 2006, 21:33:46

Job time : 185 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:30:52 ; Search time 33 Seconds
(without alignments)
460.985 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADERHKHAKDKAKSHGFIK 226

Scoring table: BIOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 239914 seqs, 67312017 residues

Total number of hits satisfying chosen parameters: 239914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications 1A New:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB.pdb:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB.pdb:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB.pdb:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB.pdb:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB.pdb:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB.pdb:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB.pdb:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB.pdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	436	36.9	843	6	US-10-449-902-47195
2	401	34.0	841	6	US-10-953-349-33825
3	375	31.8	836	7	US-11-292-431-37
4	276	23.4	784	6	US-10-953-349-33826
5	271	22.9	100	6	US-10-471-571A-3256
6	251	21.3	136	6	US-10-471-571A-3278
7	172	14.6	876	6	US-10-449-902-31438
8	172	14.6	936	6	US-10-449-902-56037
9	169	14.3	327	7	US-11-056-355B-60476
10	169	14.3	336	7	US-11-056-355B-60475
11	166	14.1	508	6	US-10-953-349-6491
12	166	14.1	508	7	US-11-056-355B-29545
13	166	14.1	508	7	US-11-056-355B-33135
14	166	14.1	508	7	US-11-056-355B-40436
15	166	14.1	508	7	US-11-056-355B-98803
16	166	14.1	508	7	US-11-056-355B-110042
17	166	14.1	595	6	US-10-953-349-6490
18	166	14.1	595	7	US-11-056-355B-29544
19	166	14.1	595	7	US-11-056-355B-33134
20	166	14.1	595	7	US-11-056-355B-40435
21	166	14.1	595	7	US-11-056-355B-98802
22	166	14.1	595	7	US-11-056-355B-110041
23	166	14.1	627	6	US-10-953-349-6489
24	166	14.1	627	7	US-11-056-355B-29543
25	166	14.1	627	7	US-11-056-355B-33133

26	166	14.1	627	7	US-11-056-355B-40434	Sequence 40434, A
27	166	14.1	627	7	US-11-056-355B-98801	Sequence 98801, A
28	166	14.1	627	7	US-11-056-355B-110040	Sequence 110040, A
29	164	13.9	419	7	US-11-056-355B-62104	Sequence 62104, A
30	164	13.9	434	7	US-11-056-355B-62103	Sequence 62103, A
31	164	13.9	464	7	US-11-056-355B-62102	Sequence 62102, A
32	163	13.8	652	7	US-11-293-697-3192	Sequence 3192, AP
33	162	13.7	817	6	US-10-449-902-51702	Sequence 51702, A
34	161	13.6	478	6	US-10-471-571A-1424	Sequence 1424, AP
35	160	13.5	656	7	US-11-056-355B-14120	Sequence 14120, A
36	160	13.5	693	7	US-11-056-355B-14119	Sequence 14119, A
37	160	13.5	722	7	US-11-056-355B-14118	Sequence 14118, A
38	159	13.5	973	6	US-10-449-902-53285	Sequence 53285, A
39	158	13.4	374	6	US-10-471-571A-4516	Sequence 51702, AP
40	158	13.4	551	7	US-11-282-262-5	Sequence 5, Appl
41	157	13.3	415	7	US-11-330-403-14269	Sequence 14269, A
42	157	13.3	473	7	US-11-330-403-5481	Sequence 5481, AP
43	157	13.3	602	7	US-11-056-355B-88014	Sequence 88014, A
44	157	13.3	604	7	US-11-056-355B-88013	Sequence 88013, A
45	157	13.3	639	7	US-11-056-355B-88012	Sequence 88012, A

ALIGNMENTS

RESULT 1

US-10-449-902-47195
; Sequence 47195, Application US/10449902
; Publication No. US20060123505A1

GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205V1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 47195
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-47195

Query Match	36.9%;	Score 436;	DB 6;	Length 843;
Best Local Similarity	37.6%;	Pred. No. 2.2e-08;		
Matches	94;	Conservative 48;	Mismatches 80;	Indels 28; Gaps 6;
QY	1	VKLTPKEQKFLLYAGEVARKKRKAEGKLNQPEAIYISAHIMDEARRKKTVAOLMEE	60	
Db	1	MKLQVREAEKLAHNAFLAQKRLARGLNVTVAALIAAQILEFVRDGRDRTVTDLMDL	60	
QY	61	CMHFLKDEVMPCGVNMVDPDLGVEATFPDGTCLVTNNWIEPDE-----H-----	105	
Db	61	GKQLLRQVLPAVPPQLLETQVQEGTFMDGTCLITVHDPISSDDGNLEALHSGFLFVPS	120	
QY	106	---PKA-----GSKFGCDKDIENAGKEVTELEVTNEGPKSLHVGSHFHFEEANKAL	155	
Db	121	LEKSSVGVDFFPGVVRP-CSGHIVLNLHRRALTALKVNKADRPITQIGHYHFIENPYL	179	
QY	156	KFDREKAYGKRLDIPSGNTRLRIGAGTKRKVQLPIPGSKKVI-GMGNLVNNTIADERHKH	214	
Db	180	VFDQRAYGMLNIPAGTAVRPEPGDAKTVTLVSIGG-RKVIRGGNGIADGAVNRSQLNE	238	
QY	215	ALDKAKSHGF	224	
Db	239	VMEKVIANGF	248	


```
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3256
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(100)
; OTHER INFORMATION: urease gamma subunit
US-10-471-571A-3256

Query Match      22.9%; Score 271; DB 6; Length 100;
Best Local Similarity 55.7%; Pred. No. 0.00048;
Matches 54; Conservative 18; Mismatches 24; Indels 1; Gaps 1;

Qy 4 TPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMESCWH 63
Db 4 TQREQDLMIVVAAEVARRRKARGLKLHHPALALISDELLEGARDG-KTVAELMSYGRQ 62

Qy 64 FLKDEVMPGVGNMVPDLGVEATPDGKLVTVANWPI 100
Db 63 ILNKEDVMDGVEHMTDIEIEATPDGKLVTVHHP 99

RESULT 6
US-10-471-571A-3278
; Sequence 3278, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471-571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3278
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(136)
; OTHER INFORMATION: urease beta subunit
US-10-471-571A-3278

Query Match      21.3%; Score 251; DB 6; Length 136;
Best Local Similarity 41.7%; Pred. No. 0.0062;
Matches 50; Conservative 21; Mismatches 42; Indels 7; Gaps 4;

Qy 109 GEV--KFGCDKDIENAGKEVTELEVNEGPKSLHVGSHFFFEANKALKPDEKAYGKR 166
Db 4 GEITK---STEVINNHHPETVTEVENTGDRPIQVGHFFFEANALDPEREMAYGKH 60

Qy 167 LDIPSGNTLIGAGQTRKVLIPLGSKKVGIMNGLVNNIADERHKHAKLDKAKSH-G-F 224
Db 61 LDIPAGAAVFEPGCKEVLVEAGRKIFGFRGMVNGPIDESRVYRPTDENDEYAGVF 120

RESULT 7
US-10-449-902-31438
; Sequence 31438, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56037
; LENGTH: 936
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-31438

Query Match      14.6%; Score 172; DB 6; Length 936;
Best Local Similarity 24.2%; Pred. No. 7.7e+02;
Matches 62; Conservative 45; Mismatches 95; Indels 54; Gaps 22;

Qy 14 YYAGEVARRRKAEGLKLNQPEAIAYISAH--I-----MDEARRGKKTVAQLMESCWHF 64
Db 516 YFSGSL--KDAAKSLGV-CPTTLKRICRHGIGSRWPSRKINKVRSILKKI-QTVINSVHG 571

Qy 65 LKKD-EVMPGVGNMVP--DLGVEATFP--DGTKLVTNNPIEDPEHFKAGEVKGCC--D 116
Db 572 VDRSLQYDPATGSLVPVWSLPEKLTFFPSCDGLPTPSVGKTVENSDLKSEE--GCSLPD 628

Qy 117 KDIELNAGKEVTELEVNEGPKSLHVGSHFFFEANKALKPDEKAYGKRLDIPSG--N 173
Db 629 GSQROSCQLQISDVKKSNB--DEFHIGSGNSDFYGANATAKGNSEVTQGP-L-CPTGAFS 684

Qy 174 TIRI-GAGQTR-KVOLIPLGGS--KKVIGMNG-----LVNNIADER-HKHKA--- 215
Db 685 ALHLKGDTCTNPSSSLRPSSESTRNQIVGRNSPSIQEQDLDMLDNHEADKDHHPSTSG 744

Qy 216 -LD----KAKSHGFTK 226
Db 745 MTDSSSGSSASHPTPK 760

RESULT 8
US-10-449-902-56037
; Sequence 56037, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56037
; LENGTH: 936
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-56037

Query Match      14.6%; Score 172; DB 6; Length 936;
Best Local Similarity 24.2%; Pred. No. 7.7e+02;
Matches 62; Conservative 45; Mismatches 95; Indels 54; Gaps 22;

Qy 14 YYAGEVARRRKAEGLKLNQPEAIAYISAH--I-----MDEARRGKKTVAQLMESCWHF 64
```

```
Db 576 YFSGSL--KDAAKSLGV-CPTTLKRICHHGISRWPSPKINKVNSLKKI-QTVINSVHG 631
Qy 65 LKXD-EVMPGVGNMVP--DLGVEATFP--DTGLVTVNWPPIEPDEHFKAGRVKFGC---D 116
Db 632 VDRSLQYDPATGSLVVPVSLPPEKLTFFPSCDGLPTFPVSGKTVBENSCLKSEE---GCSLPD 688
Qy 117 KDIELNAGKEVTELEVNEGPKSLHVG-S-HHFFPEANKALKFDEKAYGKRLDIPSG--N 173
Db 689 GSQRQSCOLQISDVKKSNB--DEFHIGSGNSDFYGANATAKSNSEVTQGP-L-CPTGAPS 744
Qy 174 TLRI-GAGQTR-KVQLIPLGGS--KKVIGMNG-----LVNNIADER-HKHKA--- 215
Db 745 ALHLKGTCTNPSSSLRPSSESTRNQIVGRNSPSIQOEDLDMLDNHEAEDKDHMPSTSG 804
Qy 216 -LD----KAKSHGFIK 226
Db 805 MTDSSSGSASSHPTFK 820

RESULT 9
US-11-056-355B-60476
; Sequence 60476, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 60476
; LENGTH: 327
; TYPE: prt
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(327)
; OTHER INFORMATION: Ceres Seq. ID no. 15169774
US-11-056-355B-60476

Query Match 14.3%; Score 169; DB 7; Length 327;
Best Local Similarity 29.3%; Pred. No. 85;
Matches 75; Conservative 32; Mismatches 88; Indels 61; Gaps 28;

Qy 1 VKLT--PKEQE--KFLYYAGEVARKRKAEGKLKNOPEAIAIYISAHIMDEA--RRGKKT 54
Db 104 VKGTTYDKVSDAIKHKLDAGNVIQ---EGLD-RAKEAI-HRATHTTQEGVQETG-NTV 156
Qy 55 AQLMEECMHFLKDE--VMPGVGNMVPDL-GV--EATFPDGTGLV--TVNWPPIEP--DEHF 106
Db 157 DETVEKVEH--RTDETVMK-VGHRITADAPGAIREGVH-DAAEAVKEAIIH-PHEPTLGERI 211
Qy 107 --KAGEVKGCDKDIELNAGKEVTELVNKGPKSLHVGSHFFHFEANKALK--FDR--E 160
Db 212 MEKINDVTHSIKAVGLEAEASIGR-KVLEKGEQK-----AASKAAKSGLEKVV 260
Qy 161 KAYGKRLDIPSGNTLRIGAGQTRKVLQIPLG-----GSKKV--IGMGLVNNIADERHKH 213
Db 261 KASGWTLETVAG--PIGAKIAEKV---AGKAASKGIEKAEQKGLNDVAGDLAEI-KE 312
Qy 214 KA--L-DKAKSHGFIK 226
Db 313 KAQHLGEKIQKHAHK 328

RESULT 11
US-10-953-349-6491
; Sequence 6491, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 6491
; LENGTH: 508
; TYPE: PRP
; ORGANISM: Arabidopsis thaliana
US-10-953-349-6491

Query Match 14.1%; Score 166; DB 6; Length 508;
Best Local Similarity 29.2%; Pred. No. 3.1e+02;
Matches 80; Conservative 33; Mismatches 71; Indels 90; Gaps 35;

Qy 8 QKFLYYAGEVAR--KRKAEGKLKNO-----PEAIA-----YI-----SAHIMDEAR 48
Db 196 QSKQLKYTGNPDQVCQTRSE-LEENQMKQYRWEQEQISHMKEYIARFPGHGSAKLARQAO 254
```

```
Db 576 YFSGSL--KDAAKSLGV-CPTTLKRICHHGISRWPSPKINKVNSLKKI-QTVINSVHG 631
Qy 65 LKXD-EVMPGVGNMVP--DLGVEATFP--DTGLVTVNWPPIEPDEHFKAGRVKFGC---D 116
Db 632 VDRSLQYDPATGSLVVPVSLPPEKLTFFPSCDGLPTFPVSGKTVBENSCLKSEE---GCSLPD 688
Qy 117 KDIELNAGKEVTELEVNEGPKSLHVG-S-HHFFPEANKALKFDEKAYGKRLDIPSG--N 173
Db 689 GSQRQSCOLQISDVKKSNB--DEFHIGSGNSDFYGANATAKSNSEVTQGP-L-CPTGAPS 744
Qy 174 TLRI-GAGQTR-KVQLIPLGGS--KKVIGMNG-----LVNNIADER-HKHKA--- 215
Db 745 ALHLKGTCTNPSSSLRPSSESTRNQIVGRNSPSIQOEDLDMLDNHEAEDKDHMPSTSG 804
Qy 216 -LD----KAKSHGFIK 226
Db 805 MTDSSSGSASSHPTFK 820

RESULT 9
US-11-056-355B-60476
; Sequence 60476, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 60476
; LENGTH: 327
; TYPE: prt
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(327)
; OTHER INFORMATION: Ceres Seq. ID no. 15169774
US-11-056-355B-60476

Query Match 14.3%; Score 169; DB 7; Length 327;
Best Local Similarity 29.3%; Pred. No. 85;
Matches 75; Conservative 32; Mismatches 88; Indels 61; Gaps 28;

Qy 1 VKLT--PKEQE--KFLYYAGEVARKRKAEGKLKNOPEAIAIYISAHIMDEA--RRGKKT 54
Db 95 VKGTTYDKVSDAIKHKLDAGNVIQ---EGLD-RAKEAI-HRATHTTQEGVQETG-NTV 147
Qy 55 AQLMEECMHFLKDE--VMPGVGNMVPDL-GV--EATFPDGTGLV--TVNWPPIEP--DEHF 106
Db 148 DETVEKVEH--RTDETVMK-VGHRITADAPGAIREGVH-DAAEAVKEAIIH-PHEPTLGERI 202
Qy 107 --KAGEVKGCDKDIELNAGKEVTELVNKGPKSLHVGSHFFHFEANKALK--FDR--E 160
Db 203 MEKINDVTHSIKAVGLEAEASIGR-KVLEKGEQK-----AASKAAKSGLEKVV 251
Qy 161 KAYGKRLDIPSGNTLRIGAGQTRKVLQIPLG-----GSKKV--IGMGLVNNIADERHKH 213
Db 252 KASGWTLETVAG--PIGAKIAEKV---AGKAASKGIEKAEQKGLNDVAGDLAEI-KE 303
Qy 214 KA--L-DKAKSHGFIK 226
Db 304 KAQHLGEKIQKHAHK 319

RESULT 10
US-11-056-355B-60475
; Sequence 60475, Application US/11056355B
; Publication No. US20060150283A1
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:24:06 ; Search time 40 Seconds
(without alignments)
543.625 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181
Sequence: 1 VKLTPKEQKFLYYAGEVA.....ADERHKHALDKAKSHGFIK 226

Scoring table: BIOSUM62
Gapop 4.0 , Gapext 1.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	619	52.4	238	2 B71977	urease alpha chain
2	616	52.2	238	1 URKCAP	urease (EC 3.5.1.5)
3	572	48.4	237	1 S35290	urease (EC 3.5.1.5)
4	511	43.3	228	2 D75586	urease, beta/gamma
5	462	39.1	840	1 URJB	urease (EC 3.5.1.5)
6	427	36.2	838	2 A96599	probable urease F1
7	416	35.2	227	2 T29055	urease (EC 3.5.1.5)
8	412	34.9	835	2 T37939	urease (EC 3.5.1.5)
9	287	24.3	100	1 A36138	urease (EC 3.5.1.5)
10	284	24.0	100	2 H70664	probable urea prot
11	283	24.0	100	2 D85603	probable urease st
12	283	24.0	100	2 B90794	urease gamma subun
13	282	23.9	100	2 S74889	urease (EC 3.5.1.5)
14	280	23.7	105	2 B69729	urease (gamma subu
15	276	23.4	100	2 A36950	urease (EC 3.5.1.5)
16	275	23.3	100	2 S47102	urease (EC 3.5.1.5)
17	274	23.2	137	1 S38484	urease (EC 3.5.1.5)
18	273	23.1	100	2 A64076	urease (EC 3.5.1.5)
19	271	22.9	100	2 C90027	urease gamma subun
20	270	22.9	100	2 B47090	urease (EC 3.5.1.5)
21	268	22.7	100	1 S08478	urease (EC 3.5.1.5)
22	267	22.6	100	1 B43719	urease (EC 3.5.1.5)
23	267	22.6	100	2 B83037	urease (beta subun
24	267	22.6	124	2 C69729	urease gamma chain
25	265	22.4	100	2 AD2264	urease (EC 3.5.1.5)
26	263	22.3	100	2 S42602	urease gamma subun
27	262	22.2	100	2 D83681	urease gamma subun
28	257	21.8	100	2 AE2872	urease gamma subun
29	257	21.8	107	2 F97648	urease gamma chain

30	257	21.8	126	2 S47103	urease (EC 3.5.1.5)
31	257	21.8	164	1 S31418	urease (EC 3.5.1.5)
32	254	21.5	108	2 AE2264	urease beta chain
33	252	21.3	100	2 S38483	urease (EC 3.5.1.5)
34	252	21.3	136	2 D90027	urease beta subuni
35	249	21.1	159	2 AB3333	urease (EC 3.5.1.5)
36	248	21.0	100	2 AI0324	urease (EC 3.5.1.5)
37	243	20.6	101	2 AC2872	urease beta subuni
38	243	20.6	102	2 S31417	urease (EC 3.5.1.5)
39	241	20.4	101	1 S42604	urease (EC 3.5.1.5)
40	238	20.2	100	2 AC3333	urease (EC 3.5.1.5)
41	238	20.2	105	2 S76224	hypothetical prote
42	237	20.1	124	2 H82890	urease complex com
43	236	20.0	101	2 A82891	urease complex com
44	236	20.0	138	2 T50709	urease (EC 3.5.1.5)
45	234	19.8	100	2 T50708	urease (EC 3.5.1.5)

ALIGNMENTS

RESULT 1

B71977

urease alpha chain - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C:Accession: B71977

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: B71977

A>Status: preliminary

A:Molecule type: DNA

A:Cross-references: UNIPROT:Q9ZM24; UNIPARC:UP10000137D6F; GB:AE001439; NID:

A:Experimental source: strain J99

C:Genetics:

A:Gene: urea

C:Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K ch

F:1-101/Domain: urease 11K chain homology <U1>

F:108-204/Domain: urease 12K chain homology <U12>

Query Match	52.4%	Score 619;	DB 2;	Length 238;
Best Local Similarity	53.1%	Pred. No. 1e-15;		
Matches 119;	Conservative 42;	Mismatches 62;	Indels 1;	Gaps 1;
Qy	1	VKLTPKEQKFLYYAGEVARKKAEGLKINQPEALAIYISAHIMDEARRGKKTVAQIMEE	60	
Db	1	MKLTPKELDKLMLHVAEGLARKEKGIKNVVEAVALISAHIMEEARRAGKKTAAELMQE	60	
Qy	61	CMHFLKDEVMGVGNMVPDLGVRATPDGTGLTVTNVPIEDPHKAGEVKFGCDKIDIE	120	
Db	61	GRTLKPDVMDGVASMIHEVGIEAMFPDGTGLTVTHTPIEANGKLVPGEL-FLKNEDIT	119	
Qy	121	LNAGKEVTEVTNEGPKSLVGHSHFHPPEANKALKFDEKAYKGLDIPSGNLTIRIGAG	180	
Db	120	INEGKAVSVKVKVNGDRFPVQIGSHFHPFEVNRCLDFDEKTFGKRLLDIASGTAVRPEP	179	
Qy	181	QTRKVLQPLPGSKKVGVMGNLVNNIADERHKHAKLDKAKSHGF	224	
Db	180	EKSVELIDIGNRRIFGFNALVDQADNESKKIALHRAKERGF	223	

RESULT 2

URKCAP

urease (EC 3.5.1.5) 26K chain - Helicobacter pylori (strains 26695 and others)

N:Alternate names: urea amidohydrolase; urease alpha chain

C:Species: Helicobacter pylori

C:Date: 30-Sep-1991 #sequence_revision 15-Aug-1997 #text_change 09-Jul-2004

C:Accession: A38537; A64529; S07884; B35306; A41502; A61371; E49215

F;1-101/Domain: urease 11K chain homology <U11>
F;108-204/Domain: urease 12K chain homology <U12>

Query Match	52.2%;	Score 616;	DB 1;	Length 238;
Best Local Similarity	52.7%;	Pred. No. 1.3e-15;		
Matches 118;	Conservative 43;	Mismatches 62;	Indels 1;	Gaps 1;

QY	1	VKLTPKQEKFLYYAGEVARKKRKAEGKLKLNQPEATAYISAHIMDBARRGKKTVAQLMEE	60
DB	1	MKLTPELKDKLMLHYAGELAKKKRKEGIKLNVVEAVALISAHIMEERAGKKTAAELMQE	60
QY	61	CHMFLKDEVMGCVGNVPDLGVEATFPDGTGLVTNNWPLEPDEHFKAGVKGCCKDIE	120
DB	61	GRTLKPDVMDGVASMIHEVGIEAMFPDGTGLVTVHTPIEANGKLVPGSEL-FLKNEDIT	119
QY	121	LNAGKEVTELEVTNEGPKSLHVGSHFPHFEANKALKFDRKAYGKRLDIPSGNLTIRIGAG	180
DB	120	INEGKAVSKVKNVGNDRPVQIGSHFHFVEVNRCLDFDRKTFGKRLDIASGTAVRPEFG	179
QY	181	QTRKVOLIPLGGSKKVIYGMNGLVNNIADRRHKHALDKAKSHGF	224
DB	180	EKSVELIDIGGNRIYGFNALVDROADNESKIALHRAKERGF	223

RESULT 3

U35290
urease (EC 3.5.1.5) 26K chain - Helicobacter felis
N:Alternate names: urease alpha chain
C:Species: Helicobacter felis
C:Date: 31-Dec-1993 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C:Accession: S35290
R:Ferro, R.L.; Labigne, A.
Mol. Microbiol. 9, 323-333, 1993
A:Title: Cloning, expression and sequencing of Helicobacter felis urease genes.
A:Reference number: S35290; MUID:94018627; PMID:8412683
A:Accession: S35290
A:Molecule type: DNA
A:Residues: 1-237 <FER>
A:Cross-references: UNIPROT:Q08715; UNIPARC:UPI0000137D6C; EMBL:X69080; NID:G399
C:Genetics:
A:Gene: ureA
C:Complex: heterodimer
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide
C:Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease
C:Keywords: heterodimer; hydrolase; metalloprotein
F;1-101/Domain: urease 11K chain homology <U11>
F;108-204/Domain: urease 12K chain homology <U12>

Query Match	48.4%;	Score 572;	DB 1;	Length 237;
Best Local Similarity	50.0%;	Pred. No. 5.8e-14;		
Matches 112;	Conservative 43;	Mismatches 68;	Indels 1;	Gaps 1;

QY	1	VKLTPKQEKFLYYAGEVARKKRKAEGKLKLNQPEATAYISAHIMDBARRGKKTVAQLMEE	60
DB	1	MKLTPELKDKLMLHYAGRLAEEALARGVKLNVTEAVALISGRVWEKARDGNKSVADLMQE	60
QY	61	CHMFLKDEVMGCVGNVPDLGVEATFPDGTGLVTNNWPLEPDEHFKAGVKGCCKDIE	120
DB	61	GRTLKKNVMDGVASMIHEVGIEAMFPDGTGLVTIHTPVEDNGKLAPGSEV-FLKNEDIT	119
QY	121	LNAGKEVTELEVTNEGPKSLHVGSHFPHFEANKALKFDRKAYGKRLDIPSGNLTIRIGAG	180
DB	120	INAGKAISLKVQNKGDPRPVQGVSHFHFVEVNRCLDFDRAKSKCKRLDIASGTAVRPEFG	179
QY	181	QTRKVOLIPLGGSKKVIYGMNGLVNNIADRRHKHALDKAKSHGF	224
DB	180	EKSVELIDIGGNRIYGFNSLVDRQADGKGLGLKRAKEKGF	223

RESULT 4

D75586
urease, beta/gamma subunit - Deinococcus radiodurans (strain R1)

A;Accession: A96699
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-838 <STO>
A;Cross-references: UNIPARC:UPI00000ABA6F; GB:AE005173; NID:G6382524; PIDN:AAF07806.1; C
C;Genetics:
A;Gene: F12B7.10
A;Map position: 1
C;Superfamily: urease; urease 11K chain homology; urease 12K chain homology; urease 62K

Query Match 36.2%; Score 427; DB 2; Length 838;
Best Local Similarity 38.0%; Pred. No. 3e-07;
Matches 100; Conservative 47; Mismatches 73; Indels 43; Gaps 9;

Qy 1 VKLTPKQEKFLYYAGEVARKKAEGKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLLPRIEXKLELHQAGFLAQKRLARGIRLNYTEAVALLIAQILEFIRDGKSVLAELMDI 60
Qy 61 CMHFLKDEVMVPGVGNMVPDLGVEATPDGKLVTVNWPI-----E----- 101
Db 61 GRQLLGRQVLPVALLHYTVQVEGTFRDGKLVTVHEPISENGNLELALHGSFLPVPS 120
Qy 102 ----PDH-----FKAGEVKFGCDKDIENAGKEVTELEVTNEGPKSLHVGSHFHFPEANKA 154
Db 121 LDKPEVEHGVIIIPGDMKIY-DGSIINHGKAVLVKVNTGDRPVQVGSYHFIEVNPL 179
Qy 155 LKFDREKAYKRLDIPSGNTLRIGAGOTRKVQLIPLGGSKKIVGMGNLVNNIAD----- 208
Db 180 LVFDRKALGNRLNIPAGTAVRFPFEGKSVLVNIGNKVIRGNGIVDGLVDDVNVTV 239
Qy 209 ----ERH--KH-KALDKAKSHG 223
Db 240 LMETMERRGFKHLEDID-A-SEG 260

RESULT 7

T29055
urease (EC 3.5.1.5) beta/gamma chain - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 06-Jan-2003
C;Accession: T29055
R;Redenbach, M.; Kieser, H.M.; Denapate, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A;Title: A set of ordered cosmid and a detailed genetic and physical map for the 8 Mb S
A;Reference number: 220556; MUID:97000351; PMID:8843436
A;Accession: T29055
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-227 <RED>
A;Cross-references: UNIPARC:UPI00000DADF4; EMBL:AL031124; PIDN:CAA19973.1
C;Genetics:
A;Gene: ureAB
C;Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K c
C;Keywords: hydrolase

Query Match 35.2%; Score 416; DB 2; Length 227;
Best Local Similarity 40.5%; Pred. No. 3.7e-08;
Matches 94; Conservative 46; Mismatches 76; Indels 16; Gaps 8;

Qy 1 VKLTPKQEKFLYYAGEVARKKAEGKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MRLTPTERDRLFLFGAELARARRARGRLNVPEATALIADTVCEARDGAR-LAQATER 59
Qy 61 CMHFLKDEVMVPGVGNMVPDLGVEATPDGKLVTVNWPIE-----PDHFKAGEVKFGCD 116
Db 60 ARSVLGPDDVLPFGVADVTVHVAVFDDGSRSLAVADPVGGGLGDD--APGALLPHGD 117
Qy 117 KDIELNAGKEVTELEVTNEG--PKSLHVGSHFHFPEANKALKFDREKAYKRLDIPSGNT 174
Db 118 RP-EPEA---ALRLPVNTATVPVS--VTSFHFPEANPRDLDGRAYGNRLAVPAGSS 171
Qy 175 LRIGAGOTRKVQLIPLGGSKKIVGMGNLVNNIAD--ERHKHAKLDKAKSHGFI 225
Db 175 LRIGAGOTRKVQLIPLGGSKKIVGMGNLVNNIAD--ERHKHAKLDKAKSHGFI 225

Db 172 VRFGPERVEVLVPIGGARVAIGFAGLVDPGADPGAREEARRAACGYL 223

RESULT 8

T37939
urease (EC 3.5.1.5) - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 12-Jul-2004
C;Accession: T37939
R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21755
A;Accession: T37939
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-835 <MCD>
A;Cross-references: UNIPROT:O00084; UNIPARC:UPI0000137DA9; EMBL:AL109820; PIDN:CAB52575.1
A;Experimental source: strain 972h-; cosmid c1952
C;Genetics:
A;Gene: SPDB:SPAC1952.11c
A;Map position: 1
C;Superfamily: urease; urease 11K chain homology; urease 12K chain homology; urease 62K c
C;Keywords: hydrolase

Query Match 34.9%; Score 412; DB 2; Length 835;
Best Local Similarity 37.4%; Pred. No. 1.1e-06;
Matches 95; Conservative 39; Mismatches 87; Indels 33; Gaps 7;

Qy 3 LTPKQEKFLYYAGEVARKKAEGKLNQPEAIAVISAHIMDEARRGKKTVAQLMEECM 62
Db 1 MOPRELHLTLHLQGLSLAQKRLCRGKLNKLEATSLIASQIQEYVRDGNHNSVADLSLKG 60
Qy 63 HFLKXDEVMPGVGNMVPDLGVEATPDGKLVTVNWPI-EPD---EH----- 105
Db 61 DMLGKHVQPNVLLHHEIMIEATPDGTYLTIHDPICTDGNLEHALYGSFLPTPSQE 120
Qy 106 -FKAGEVKF-----G-----CDKDIEL--NAGKEVTELEVTNEGPKSLHVGSHFHFPEA 151
Db 121 LPLPEEKLYAPENSPGFVELEGTELLPNLPR--TPIEVNMGDRPIQVGSYHFIEIET 178
Qy 152 NKALFDEKAYKRLDIPSGNTLRIGAGOTRKVQLIPLGGSKKIVGMGNLVNNIADDERH 211
Db 179 NEKLCFDRSKAYKRLDIPSGTAIRPEPGVMKIVNLIPIGGAKLIQGNLSKGVFDDSR 238
Qy 212 KHKALDKAKSHGFI 225
Db 239 TREIVDNLKQSGFM 252

RESULT 9

A36138
urease (EC 3.5.1.5) 11K chain [validated] - Klebsiella pneumoniae
N;Alternate names: urea amidohydrolase; urease chain A; urease gamma chain
C;Species: Klebsiella pneumoniae
C;Date: 30-Nov-1990 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: A36138; S32938
R;Mulrooney, S.B.; Hausinger, R.P.
J. Bacteriol. 172, 5837-5843, 1990
A;Title: Sequence of the Klebsiella aerogenes urease genes and evidence for accessory pr
A;Reference number: A36138; MUID:91008957; PMID:2211515
A;Accession: A36138
A;Molecule type: DNA
A;Residues: 1-100 <MUL>
A;Cross-references: UNIPROT:P18316; UNIPARC:UPI00001103C0; GB:M36068; NID:g149335; PIDN:?
R;Collins, C.M.; Gutman, D.M.; Laman, H.
Mol. Microbiol. 8, 187-198, 1993
A;Title: Identification of a nitrogen-regulated promoter controlling expression of Klebsi
A;Reference number: S32937; MUID:93268094; PMID:8497192
A;Accession: S32938
A;Molecule type: DNA
A;Residues: 1-66, 'K', 68-100 <COL>
A;Cross-references: UNIPARC:UPI0000137D9A; EMBL:L07039; NID:g149330; PIDN:AAA25147.1; PII
R;Jabri, E.; Karplus, P.A.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:20:41 ; Search time 300 Seconds

(without alignments)
696.845 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLYYAGEVA.....ADERHKHAKDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	665	56.3	234	1 UREA_HELBI	Q8q98 helicobacte
2	638	54.0	234	1 UREA_HELHE	P42822 helicobacte
3	619	52.4	238	1 UREA_HELJP	Q9zm24 helicobacte
4	617	52.2	238	2 Q8RNT5_HELVP	Q8rmu5 helicobacte
5	617	52.2	238	2 Q93N11_HELVP	Q93nj1 helicobacte
6	616	52.2	238	1 UREA_HELJP	P14916 helicobacte
7	616	52.2	238	2 Q93N13_HELVP	Q93ni3 helicobacte
8	616	52.2	238	2 Q93N15_HELVP	Q93ni5 helicobacte
9	615	52.1	238	2 Q93N18_HELVP	Q93ni8 helicobacte
10	614	52.0	238	2 Q933M9_HELVP	Q933m9 helicobacte
11	613	51.9	238	2 Q932S5_HELVP	Q932s5 helicobacte
12	613	51.9	238	2 Q93N14_HELVP	Q93ni4 helicobacte
13	613	51.9	238	2 Q93N10_HELVP	Q93nj0 helicobacte
14	613	51.9	238	2 Q93N12_HELVP	Q93nj2 helicobacte
15	611	51.7	238	2 Q84P76_HELVP	Q84f76 helicobacte
16	610	51.7	238	2 Q93N19_HELVP	Q93ni9 helicobacte
17	609	51.6	238	2 Q93NH7_HELVP	Q93nh7 helicobacte
18	609	51.6	238	2 Q93N12_HELVP	Q93ni2 helicobacte
19	608	51.5	238	2 Q933P5_HELVP	Q933p5 helicobacte
20	607	51.4	238	2 Q93N10_HELVP	Q93ni0 helicobacte
21	607	51.4	238	2 Q93N11_HELVP	Q93ni1 helicobacte
22	605	51.2	238	2 Q93NH8_HELVP	Q93nh8 helicobacte
23	605	51.2	238	2 Q93N16_HELVP	Q93ni6 helicobacte
24	604	51.1	238	2 Q93N17_HELVP	Q93ni7 helicobacte
25	602	51.0	238	2 Q93NH9_HELVP	Q93nh9 helicobacte
26	572	48.4	237	1 UREA_HELFE	Q8w715 helicobacte
27	561	47.5	347	2 Q7PED9_ANOGA	Q7ped9 anopheles g
28	556	47.1	225	1 UREA_HELHP	Q93pj5 helicobacte
29	554	46.9	213	1 UREA_HELHU	P50044 helicobacte
30	529	44.8	219	1 URE23_SULTO	Q972v9 sulfolobus
31	511	43.3	228	1 URE23_DBIRA	Q9ryj3 deinococcus

32	505	42.8	231	1 URE23_PSESM	O883f3 pseudomonas
33	503	42.6	223	2 Q5FB24_CAMLA	O5fb24 campylobact
34	502	42.5	231	2 Q4ZUD1_PSEU2	Q4zud1 pseudomonas
35	484	41.0	231	2 Q35P06_9BRAD	Q35p06 bradyrhizob
36	467	39.5	838	2 Q7XAC5_SOVEN	Q7xacs glycine max
37	462	39.1	840	1 UREA_CANEN	P07374 canavalia e
38	449	38.0	176	2 Q8KIY9_HELHE	O8kiy9 helicobacte
39	448	37.9	177	2 Q8KJ05_HELHE	O8kj05 helicobacte
40	446	37.8	172	2 Q8KHJ0_HELHE	O8khj0 helicobacte
41	446	37.8	179	2 Q8KI21_HELHE	Q8ki21 helicobacte
42	445	37.7	185	2 Q9R8R4_HELHP	Q9r8r2 helicobacte
43	443	37.5	185	2 Q9R8R2_HELHP	Q9r8r2 helicobacte
44	439	37.2	164	2 Q8KIZ9_HELBI	O8kiz9 helicobacte
45	437	37.0	190	2 Q61616_CAMLA	Q61616 campylobact

ALIGNMENTS

RESULT 1
UREA_HELBI STANDARD; PRT; 234 AA.
AC Q8GH98;
DT 16-JAN-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-MAR-2003, sequence version 1.
DT 21-FEB-2006, entry version 20.
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase alpha subunit).
GN Name=ureA;
OS Helicobacter bizzozeronii.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=56877;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=2540228; PubMed=12652903; DOI=10.1080/1042517021000039230;
RA Zhu J., Teng C.H., Chang C.P., Chang C.D., Simpson K.W., Wei C.,
RA McDonough P., McDonough S., Chang Y.F.;
RT "Cloning and characterization of a Helicobacter bizzozeronii urease
gene cluster.";
RL DNA Seq. 13:321-331(2002).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit (By
similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- SIMILARITY: In the N-terminal section; belongs to the urease gamma
subunit family.
CC -!- SIMILARITY: In the C-terminal section; belongs to the urease beta
subunit family.
CC -!- CAUTION: In Helicobacter the alpha subunit is what is known, in
other bacteria as the beta subunit.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AF330621; AA015373.1; -; Genomic DNA.
HSSP; P14916; 1B9Z.
SMR; Q8GH98; 1-224.
HMAP; MF_00739; fused; 1.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRPFAMs; TIGR00192; urease_beta; 1.
DR TIGRPFAMs; TIGR00193; urease_gamma; 1.
DR Hydrolase.
KW CHAIN 1 234 Urease alpha subunit.
FT REGION 1 102 /FTID=PRO_0000098069.
FT REGION 103 234 Urease_gamma.
FT REGION 103 234 Urease_beta.

```

/FTId=PRO_0000098071.
FT REGION 1 102
FT REGION 103 234
FT REGION 103 234
SQ SEQUENCE 234 AA; 25746 MW; CD5889E3D789C7F8 CRC64;

Query Match 54.0%; Score 638; DB 1; Length 234;
Best Local Similarity 54.8%; Pred. No. 2.2e-16;
Matches 126; Conservative 41; Mismatches 55; Indels 8; Gaps 3;

QY 1 VKLTPEQEKFLYYAGEVARKEKAEGLKLNQPEAIAYISAHIMDEARRGKTKTVAOLMEE 60
DB 1 MKLTPELDELKMLHYAGELAKQKAKGKILNYTEAVALLSAHWEEARAGKKSVDLMQE 60
QY 61 CMHFLPKDDEYMPGVGNMVPDLGVEATPPDGTGKLTIVNWPTEP-DEHFKAGEV--KFCDK 117
DB 61 GRTLLKADDVMPGVAMHIEHVGIEAGPDGDKLTIVHTPTVEAGSDKLAPGEVILK--NE 117
QY 118 DIENAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDISGNTLRI 177
DB 118 DITLNAGKHAVALQVKNKGDRPVQVGSFHFFFEVKNLLDFDREKAYGKRLDIASGTAVRF 177
QY 178 GAGOTRKVOLIPLUGGSKVTGMGNLVNNIADERHGHKALKDCAK-SH-GFI 225
DB 178 EPGEETVELIDGGNKRIYGFNALVDROADHDGKQLAKRAKEKHFTI 227

RESULT 3
UREA_HELPJ STANDARD; PRT; 238 AA.
AC Q9ZMZ4;
DT 30-MAY-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-MAY-1999, sequence version 1.
DT 07-MAR-2006, entry version 37.
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase alpha subunit).
GN Name=urea; Synonyms=hpaA; OrderedLocusNames=JHP0068;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OC NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., DeJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori."
RL Nature 397:176-180 (1999).
CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -1- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit (By
similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -1- SIMILARITY: In the N-terminal section; belongs to the urease gamma
subunit family.
CC -1- SIMILARITY: In the C-terminal section; belongs to the urease beta
subunit family.
CC -1- CAUTION: In Helicobacter the alpha subunit is what is known, in
other bacteria as the beta subunit.
-----
Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
Distributed under the Creative Commons Attribution-NoDerivs License
-----
CC EMBL; AE001446; AAD05652.1; -; Genomic_DNA.
DR PIR; B71977; B71977.
DR HSSP; P14916; IE9Z.
DR SMR; Q9ZMZ4; 1-238.
DR GenomeReviews; AE001439 GR; JHP0068.
DR BioCyc; HPY185963:JHP0068-MONOMER; -.
DR HAMAP; MF 00739; fused; 1.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.

```

```
DR InterPro; IPR008223; Urease gamma/beta.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease gamma/beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
KW Complete proteome; Hydrolase.
FT CHAIN 1 238 Urease alpha subunit.
FT /FTID=PRO_000098074.
FT REGION 1 102 Urease gamma.
FT REGION 103 238 Urease_beta.
FT SEQUENCE 238 AA; 26568 MW; A10B9DC4156C0561 CRC64;
Query Match 52.4%; Score 619; DB 1; Length 238;
Best Local Similarity 53.1%; Pred. No. 1.2e-15;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDKMLHYAGELAKRKEGKIKLNYEAVALLISAHIMEEARAGKKTAAELMQE 60
QY 61 CMHFLKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNVNPIEPDEHFKAGVKGCDKDIE 120
DB 61 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTVTHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKFDREKAYGKRLDIPSGNTLRICAG 180
DB 120 INEGKAVSVKVNKVGDRPVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRPEFG 179
QY 181 QTRKVQLIPLGGSKKVGIMGLVNNIADERHKHAKLDKAKSHGF 224
DB 180 EEKSVELIDIGNRRIFGFNALVDRQADNESKIALHRAKERGF 223
RESULT 4
Q93NJ1_HELPY
ID Q93NJ1_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ1;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Urease A subunit.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1;
RX MEDLINE=20123532; PubMed=10660136;
RA Opazo P., Muller I., Rollan A., Valenzuela P., Yudelevich A.,
RA Garcia-de la Guardia R., Urra S., Venegas A.;
RA "Serological response to Helicobacter pylori recombinant antigens in
RT Chilean infected patients with duodenal ulcer, non-ulcer dyspepsia and
RT gastric cancer.";
RL APMIS 107:1069-1078 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1;
RX MEDLINE=22121498; PubMed=12125208;
RA Muller I., Medina-Selby A., Palacios J.L., Martinez P., Opazo P.,
RA Bruce E., Mancilla M., Valenzuela P., Yudelevich A., Venegas A.;
RA "Cloning and comparison of ten gene sequences of a Chilean H. pylori
RT strain with other H. pylori strains revealed higher variability for
RT VacA and CagA virulence factors.";
RL Biol. Res. 35:67-84 (2002).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR InterPro; IPR008223; Urease gamma/beta.
DR HSSP; P14916; 1E9Y.
DR SMR; Q93NJ1; 1-238.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma/beta.
DR Pfam; PF00699; Urease_beta; 1.
DR PIRSF; PIRSF001225; Urease gamma/beta; 1.
DR ProDom; PD002326; Urease_gamma; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26710 MW; E934A6512A559B28 CRC64;
Query Match 52.2%; Score 617; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 1.5e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDKMLHYAGELAKRKEGKIKLNYEAVALLISAHIMEEARAGKKTAAELMQE 60
QY 61 CMHFLKDEVMGPGVGNMVPDLGVEATFPDGTGLVTNVNPIEPDEHFKAGVKGCDKDIE 120
DB 61 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTVTHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKFDREKAYGKRLDIPSGNTLRICAG 180
DB 120 INEGKAVSVKVNKVGDRPVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRPEFG 179
QY 181 QTRKVQLIPLGGSKKVGIMGLVNNIADERHKHAKLDKAKSHGF 224
DB 180 EEKSVELIDIGNRRIFGFNALVDRQADNESKIALHRAKERGF 223
RESULT 5
Q93NJ1_HELPY
ID Q93NJ1_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ1;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein UreA.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=93k;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RA "Rate of Helicobacter pylori infection in children and clonality of
RT Taiwan strains.";
RL Microbiol. Immunol. 47:813-821 (2003).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AF373558; AAKG9724.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SMR; Q93NJ1; 1-238.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
```

```
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_gamma; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26672 MW; EE82BF861D03BA50 CRC64;

Query Match 52.2%; Score 617; DB 2; Length 238;
Best Local Similarity 53.1%; Pred. No. 1.5e-15;
Matches 119; Conservative 43; Mismatches 61; Indels 1; Gaps 1;

Qy 1 VKLTPKQEQELLYAGVAVKRAEGLKNOPRAIAVISAHIMDEARRGKKTVAQLMEE 60
Db :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MKLTPKELDKMLHYAGELAKRKKEGKLVNVAVALISAHIMDEARRGKKSAAELMQE 60

Qy 61 CMHFLKXDEVMPGVGNMVPDGLGVEATFPDGTGLVTVMPIEPDEHFHFKAGEVKFGCDKIE 120
Db :||| : : : : : : : : : : : : : : : : : : : : : : : : : :
61 GRTELPDDVMDGVASMIHEVGIEAMPDPGKLVTVHTFIEVNGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELVTEGPKSLHVGSHFFHFEANKALKPDREKAYKRLDIPSGNTLRIGAG 180
Db :||| : : : : : : : : : : : : : : : : : : : : : : : : : :
120 INEGKXAVSVKRVNGDPRPVQIGSHFFHFEVNYLDFDREKTFGKRLDIAAGTTVRFEPG 179

Qy 181 QTRVQVLIPLGSKKVTGMNGLVNNIADERHKHAKLDKAKSHGF 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
180 EKSVELIDIGNRRIFGFNALVDROADNESKKIALHRAKRGF 223

RESULT 6
UREA_HELPY STANDARD; PRT; 238 AA.
AC P14916;
DT 01-APR-1990, integrated into UniProtKB/Swiss-Prot.
DT 01-MAY-1992, sequence version 2.
DT 07-MAR-2006, entry version 57.
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase alpha subunit).
GN Name=urea; Synonyms=npuA; OrderedLocustNames=HP0073;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=85P;
RX MEDLINE=91161505; PubMed=2001995;
RA Labigne A., Cussac V., Courcoux P.;
RT "Shuttle cloning and nucleotide sequences of Helicobacter pylori genes responsible for urease activity.";
RL J. Bacteriol. 173:1920-1931(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=CPM630;
RX MEDLINE=90221820; PubMed=2326167;
RA Clayton C.L., Pallen M.J., Kleanthous H., Wren B.W., Tabaqchali S.;
RT "Nucleotide sequence of two genes from Helicobacter pylori encoding for urease subunits.";
RL Nucleic Acids Res. 18:362-362(1990).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Gill S.R.,
RA Fleischman R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,
RA Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,
RA Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,
RA Glodek A., McKenney K., FitzGerald L.M., Lee N., Adams M.D.,
RA Hickey E.K., Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D.,
RA Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,
RA Matthey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,
```

```
RA Smith H.O., Fraser C.M., Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter pylori.";
RL Nature 388:539-547(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=HPK5;
RA Akada J.K., Shirai M., Takeuchi H., Tsuda M., Nakazawa T.;
RT "The urease operon in Helicobacter pylori is regulated by decay of mRNA.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP PROTEIN SEQUENCE OF 1-20.
RX MEDLINE=90264448; PubMed=2188975;
RA Dunn B.E., Campbell G.P., Perez-Perez G.I., Blaser M.J.;
RT "Purification and characterization of urease from Helicobacter pylori.";
RL J. Biol. Chem. 265:9464-9469(1990).
RN [6]
RP PROTEIN SEQUENCE OF 1-20.
RX MEDLINE=90202165; PubMed=2318539;
RA Hu L.T., Mobley H.L.T.;
RT "Purification and N-terminal analysis of urease from Helicobacter pylori.";
RL Infect. Immun. 58:992-998(1990).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS), AND SUBUNIT STRUCTURE.
RX MEDLINE=21267477; PubMed=11373617; DOI=10.1038/88563;
RA Ha N.-C., Oh S.-T., Sung J.Y., Cha K.A., Lee M.H., Oh B.-H.;
RT "Supramolecular assembly and acid resistance of Helicobacter pylori urease.";
RL Nat. Struct. Biol. 8:505-509(2001).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- SUBUNIT: Heterohexamer of 3 alpha and 3 beta subunits. Four heterohexamers assemble to form a 16 nm spherical complex.
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -!- MISCELLANEOUS: Oligomerization may protect the enzyme against denaturation in an acidic environment.
CC -!- SIMILARITY: In the N-terminal section; belongs to the urease gamma subunit family.
CC -!- SIMILARITY: In the C-terminal section; belongs to the urease beta subunit family.
CC -!- CAUTION: In Helicobacter the alpha subunit is what is known, in other bacteria as the beta subunit.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; X17079; CAA34932.1; -; Genomic DNA.
EMBL; M60398; AAA25020.1; -; Genomic DNA.
EMBL; A08816; CAA00810.1; -; Unassigned DNA.
EMBL; A07396; CAA00662.1; -; Unassigned DNA.
EMBL; AB000511; AAD07144.1; -; Genomic DNA.
EMBL; AB032429; BAA84532.1; -; Genomic DNA.
PIR; A38537; URKCAP.
PDB; 1E9Y; X-ray; A=1-238.
PDB; 1E9Z; X-ray; A=1-238.
GenomeReviews; AB000511_GR; HP0073.
TIGR; HP0073; -.
HAMAP; MF_00739; fused; 1.
InterPro; IPR002019; Urease_beta.
InterPro; IPR002026; Urease_gamma_reg.
InterPro; IPR008223; Urease_gamma_beta.
Pfam; PF00699; Urease_beta; 1.
Pfam; PF00547; Urease_gamma; 1.
PIRSF; PIRSF001225; Urease_gamma_beta; 1.
ProDom; PD002326; Urease_beta; 1.
ProDom; PD002319; Urease_gamma; 1.
TIGRFAMs; TIGR00192; urease_beta; 1.
TIGRFAMs; TIGR00193; urease_gamma; 1.
3D-structure; Complete proteome; Direct protein sequencing; Hydrolase.
CHAIN 1 238 /FTId=PRO_0000098075.
FT
```



```

RESULT 8
Q93N15_HELPY
Q93N15_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93N15;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=47;
RC MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RT Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AF373568; AK69734.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SMR; Q93N15; 1-238.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016151; F.nickel ion binding; IEA.
DR GO; GO:0009039; F.urease activity; IEA.
DR GO; GO:0006807; P.nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRfams; TIGR00192; urease_beta; 1.
DR TIGRfams; TIGR00193; urease_gamma; 1.
DR SEQUENCE 238 AA; 26540 MW; 251AD890042A4262 CRC64;
SQ
Query Match 52.2%; Score 616; DB 2; Length 238;
Best Local Similarity 52.7%; Pred No. 1.6e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPQEKEFLYAYAGEVARAKRKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKLHYAGELARKKEGKILNYYEAVALISAHIMEEARAGKKSAAELMQE 60
QY 61 CMHFLKDEVMVGVMVDPGLVEATPDGTKLVTVMWPIEDPHFKAGEVKFCDDKIE 120
Db 61 GRSLKPPDDVMDGVASMIHEVGIEAMPDGDGKLVTHPTIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAKGKVEELVETNEGPKSLHVGSHFFHFFRANKALKFDREKAYGKRLDIPSGNTLRTGAG 180
Db 120 INEGKAVSVKXNVGDRPVQIGSHFFHFFVNRCLDFDREKTPGKRLDIASGTVRFPFG 179
QY 181 QTRKVLQIPLGSKKVTGMNGLVNNIADERHKHKALDKAKSHGF 224
Db 180 EEEKSVELIDGGNRRIFGFGNALVDRQADNESKKTALHRAKERGF 223

RESULT 9
Q93N18_HELPY
Q93N18_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93N18;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=47;
RC MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RT Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AF373568; AK69734.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SMR; Q93N15; 1-238.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016151; F.nickel ion binding; IEA.
DR GO; GO:0009039; F.urease activity; IEA.
DR GO; GO:0006807; P.nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRfams; TIGR00192; urease_beta; 1.
DR TIGRfams; TIGR00193; urease_gamma; 1.
DR SEQUENCE 238 AA; 26540 MW; 251AD890042A4262 CRC64;
SQ
Query Match 52.2%; Score 616; DB 2; Length 238;
Best Local Similarity 52.7%; Pred No. 1.6e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPQEKEFLYAYAGEVARAKRKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKLHYAGELARKKEGKILNYYEAVALISAHIMEEARAGKKSAAELMQE 60
QY 61 CMHFLKDEVMVGVMVDPGLVEATPDGTKLVTVMWPIEDPHFKAGEVKFCDDKIE 120
Db 61 GRSLKPPDDVMDGVASMIHEVGIEAMPDGDGKLVTHPTIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAKGKVEELVETNEGPKSLHVGSHFFHFFRANKALKFDREKAYGKRLDIPSGNTLRTGAG 180
Db 120 INEGKAVSVKXNVGDRPVQIGSHFFHFFVNRCLDFDREKTPGKRLDIASGTVRFPFG 179
QY 181 QTRKVLQIPLGSKKVTGMNGLVNNIADERHKHKALDKAKSHGF 224
Db 180 EEEKSVELIDGGNRRIFGFGNALVDRQADNESKKTALHRAKERGF 223

```

[illegible]


```
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF373580; AAK69746.1; -; Genomic_DNA.
DR EMBL; AF373572; AAK69738.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR00219; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_gamma_beta; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26600 MW; 64A0C32A043A88D8 CRC64;

Query Match 52.08; Score 614; DB 2; Length 238;
Best Local Similarity 52.78; Pred. No. 2e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPEQKFLLYAGVAVKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPELDKMLHYAGELARKKEKIGKLNLYEAVALLSAHIMEARAGKSAELMQE 60

QY 61 CMHFLKDEVMGVMGNNVDPDLGVEATPDGTKLVTVNWPIDPDEHFKAGEVKFGCDKIE 120
Db 61 GRSLLKPDVMDGVASMIHEVGIEAMFPDGTGLTVHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFHPPFANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKVNVDPRVQIGSHFFHPPFVNRVYLDREKTFGKRLDIASGTAVRPEPG 179

QY 181 QTRKVLQIPLGGSKKVGNNLVNNIADERHKHKLDAKASHGF 224
Db 180 EKSVELIDIGNRRIFGNALVDRQADNESKKIALHRAKERGF 223

RESULT 11
Q932S5_HELPY
ID Q932S5_HELPY PRELIMINARY; PRT; 238 AA.
AC Q932S5;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=100;
RX MEDLINE=100, 101, 5c, 85K, 86K, and 89K;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
```

```
CC EMBL; AF373584; AAK69750.1; -; Genomic_DNA.
DR EMBL; AF373557; AAK69723.1; -; Genomic_DNA.
DR EMBL; AF373559; AAK69725.1; -; Genomic_DNA.
DR EMBL; AF373581; AAK69747.1; -; Genomic_DNA.
DR EMBL; AF373582; AAK69748.1; -; Genomic_DNA.
DR EMBL; AF373583; AAK69749.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_gamma_beta; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26614 MW; E4A2917D016FD8DB CRC64;

Query Match 51.9%; Score 613; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 2.1e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPEQKFLLYAGVAVKRAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPELDKMLHYAGELARKKEKIGKLNLYEAVALLSAHIMEARAGKSAELMQE 60

QY 61 CMHFLKDEVMGVMGNNVDPDLGVEATPDGTKLVTVNWPIDPDEHFKAGEVKFGCDKIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGLTVHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFHPPFANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKVNVDPRVQIGSHFFHPPFVNRVYLDREKTFGKRLDIASGTAVRPEPG 179

QY 181 QTRKVLQIPLGGSKKVGNNLVNNIADERHKHKLDAKASHGF 224
Db 180 EKSVELIDIGNRRIFGNALVDRQADNESKKIALHRAKERGF 223

RESULT 12
Q93NI4_HELPY
ID Q93NI4_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NI4;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=48;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RL Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF373569; AAK69735.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SMR; Q93NI4; 1-238.
DR
```

```
DR GO:0005737; C:cytoplasm; IEA.
DR GO:0016151; F:nickel ion binding; IEA.
DR GO:0009039; F:urease activity; IEA.
DR GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro: IPR002019; Urease_beta.
DR InterPro: IPR002026; Urease_gamma_reg.
DR InterPro: IPR008223; Urease_gamma_beta.
DR Pfam: PF00699; Urease_beta; 1.
DR Pfam: PF00547; Urease_gamma; 1.
DR PIRSF: PIRSF001225; Urease_gamma_beta; 1.
DR ProDom: PD002326; Urease_gamma_beta; 1.
DR ProDom: PD002319; Urease_gamma; 1.
DR TIGRFAMs: TIGR00192; urease_beta; 1.
DR TIGRFAMs: TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26512 MW; CA6677D2788BDD2E CRC64;

Query Match          51.9%; Score 613; DB 2; Length 238;
Best Local Similarity 52.2%; Pred. No. 2.1e-15;
Matches 117; Conservative 44; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQERFLYYAGEVARKRKAEGKLKNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDKLMLHYAGELAKRKEKGKLNVEAVALISAHIMEEARAGKKSAAELMQE 60

QY 61 CMHFLKDEVMGPNVMDLGVATFPDGTGLVTNNWPIEDPDEHFKAGVKGCDKDIE 120
DB 61 GRSLLKPDVMDGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTEGPKSLHVGSHFFRANKALKDFREKAYGKRLDIPSGNLTIRIGAG 180
DB 120 INEGKKAVSVKVNVDGDRPVQIGSHFFRANKALKDFREKAYGKRLDIPSGNLTIRIGAG 179

QY 181 QTRKVLQIPLOGSKKVGNGLVNNIADRHKHKALDKAKSHGF 224
DB 180 BEKSVELIDIGNRRIFGFNALVDRQADNESKIALHRAKERGF 223

RESULT 13
Q93NJ0_HELPY
ID Q93NJ0_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ0;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=17c;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RT Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AF373560; AAK69726.1; -; Genomic_DNA.
DR HSSP: PI4916; IE9Y.
DR SMR: Q93NJ0; 1-238.
DR GO:0005737; C:cytoplasm; IEA.
DR GO:0016151; F:nickel ion binding; IEA.
DR GO:0009039; F:urease activity; IEA.
DR GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro: IPR002019; Urease_beta.
DR InterPro: IPR002026; Urease_gamma_reg.
DR InterPro: IPR008223; Urease_gamma_beta.
DR Pfam: PF00699; Urease_beta; 1.
DR Pfam: PF00547; Urease_gamma; 1.
DR ProDom: PD002326; Urease_gamma_beta; 1.
DR ProDom: PD002319; Urease_gamma; 1.
DR TIGRFAMs: TIGR00192; urease_beta; 1.
DR TIGRFAMs: TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26630 MW; E4A2917D17B5D8DB CRC64;
```

```
DR Pfam: PF00547; Urease_gamma; 1.
DR PIRSF: PIRSF001225; Urease_gamma_beta; 1.
DR ProDom: PD002326; Urease_beta; 1.
DR ProDom: PD002319; Urease_gamma; 1.
DR TIGRFAMs: TIGR00192; urease_beta; 1.
DR TIGRFAMs: TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26630 MW; 5C082660AAC88469 CRC64;

Query Match          51.9%; Score 613; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 2.1e-15;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQERFLYYAGEVARKRKAEGKLKNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 MNLTPKELDKLMLHYAGELAKRKEKGKLNVEAVALISAHIMEEARAGKKSAAELMQE 60

QY 61 CMHFLKDEVMGPNVMDLGVATFPDGTGLVTNNWPIEDPDEHFKAGVKGCDKDIE 120
DB 61 GRTLKPDVMDGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTEGPKSLHVGSHFFRANKALKDFREKAYGKRLDIPSGNLTIRIGAG 180
DB 120 INEGKKAVSVKVNVDGDRPVQIGSHFFRANKALKDFREKAYGKRLDIPSGNLTIRIGAG 179

QY 181 QTRKVLQIPLOGSKKVGNGLVNNIADRHKHKALDKAKSHGF 224
DB 180 BEKSVELIDIGNRRIFGFNALVDRQADNESKIALHRAKERGF 223

RESULT 14
Q93NJ2_HELPY
ID Q93NJ2_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ2;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease protein Urea.
GN Name=urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=96C;
RX MEDLINE=23001021; PubMed=14638992;
RA Chu C., Yu Y.J., Kong M.S., Ou J.T.;
RT "Rate of Helicobacter pylori infection in children and clonality of
RT Taiwan strains.";
RL Microbiol. Immunol. 47:813-821(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AF373556; AAK69722.1; -; Genomic_DNA.
DR HSSP: PI4916; IE9Y.
DR SMR: Q93NJ2; 1-238.
DR GO:0005737; C:cytoplasm; IEA.
DR GO:0016151; F:nickel ion binding; IEA.
DR GO:0009039; F:urease activity; IEA.
DR GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro: IPR002019; Urease_beta.
DR InterPro: IPR002026; Urease_gamma_reg.
DR InterPro: IPR008223; Urease_gamma_beta.
DR Pfam: PF00699; Urease_beta; 1.
DR Pfam: PF00547; Urease_gamma; 1.
DR ProDom: PD002326; Urease_gamma_beta; 1.
DR ProDom: PD002319; Urease_gamma; 1.
DR TIGRFAMs: TIGR00192; urease_beta; 1.
DR TIGRFAMs: TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26630 MW; E4A2917D17B5D8DB CRC64;
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:34:02 ; Search time 196 Seconds
(without alignments)
1324.995 Million cell updates/sec

Title: US-09-904-994b-3

Perfect score: 2999

Sequence: 1 MKMKKQEVNTYGTGDKV.....KLCTSKPTSQVLAQRVYTFP 568

Scoring table:

BLOSUM62
Gapop 4.0 , Gapext 1.0

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_8.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*
- 10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2999	100.0	568	5	ADJ58239 Urease su
2	2973	99.1	568	5	Adj58251 UreaseXY
3	2960	98.7	568	5	ADJ58242 UreaseXY
4	2953	98.5	568	5	ADJ58248 UreaseXY
5	2591	86.4	496	5	ADJ58245 UreaseXY
6	2326	77.6	569	8	ADQ37851 H. bizzoz
7	2297	76.6	569	2	AAR04580 Part of p
8	2297	76.6	569	2	AAR07194 H. pylori
9	2297	76.6	569	2	AAR09511 H. pylori
10	2297	76.6	569	6	ABU30687 Protein e
11	2297	76.6	569	8	ADU05536 H. pylori
12	2282	76.1	568	8	ADS09177 H. pylori
13	2282	76.1	570	4	ABE55116 Hylcobac
14	2281	76.1	569	2	AAR67378 H. pylori
15	2244	74.8	569	2	AAR74337 Helicobac
16	2244	74.8	569	2	AAR06730 H. felis
17	2243	74.8	749	2	AAR05099 Helicobac
18	2238	74.6	806	2	AAR67371 Urease A
19	2237	74.6	559	2	AAR13550 B subunit
20	2229	74.3	559	8	ADM28644 Helicobac
21	2029	67.7	568	2	AAR36387 Urease ga
22	1927	64.3	568	6	ABU19714 Protein e
23	1921	64.1	566	6	ABU41362 Protein e

24	1913	63.8	568	6	ABU21842	Abu21842 Protein e
25	1902	63.4	568	6	ABU22314	Abu22314 Protein e
26	1843	61.5	572	6	ABU30286	Abu30286 Protein e
27	1838	61.3	566	6	ABU38819	Abu38819 Protein e
28	1834	61.2	566	7	ABO78141	AbO78141 Pseudomon
29	1834	61.2	567	6	ABU32083	Abu32083 Protein e
30	1831	61.1	569	7	ADF05744	Adf05744 Bacterial
31	1831	61.1	779	6	ABU41210	Abu41210 Protein e
32	1830	61.0	625	7	ABO62699	ABO62699 Klebsiell
33	1829	61.0	567	6	ADA34625	Ada34625 Acinetoba
34	1829	61.0	568	2	AAR67379	Aar67379 P. mirabi
35	1826	60.9	566	6	ABU17040	Abu17040 Protein e
36	1826	60.9	567	2	AAW37774	Aaw37774 Klebsiell
37	1826	60.9	567	3	AAW81825	Aay81825 Klebsiell
38	1818	60.6	568	6	ABU28899	Abu28899 Protein e
39	1818	60.6	568	6	ABU28920	Abu28920 Protein e
40	1818	60.6	568	7	ADC00113	Adc00113 Enterohae
41	1810	60.4	567	6	ABU40328	Abu40328 Protein e
42	1804	60.2	570	2	AAR67380	Aar67380 Jack bean
43	1804	60.2	571	4	AAG81512	Aag81512 S. epider
44	1804	60.2	811	6	ABU42650	Abu42650 Protein e
45	1803	60.1	573	5	ABP40181	Abp40181 Staphyloc

ALIGNMENTS

RESULT 1
ADJ58239
ID ADJ58239 standard; protein; 568 AA.
XX
AC ADJ58239;
XX
DT 06-MAY-2004 (first entry)
XX
DE Urease subunit polypeptide Y.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
N-PSDB; ADJ58237.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
PT Claim 10; SEQ ID NO 3; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
XX
CC Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an urease
XX
CC Y subunit polypeptide of the invention.
XX
SQ Sequence 568 AA;

Query Match 100.0%; Score 2999; DB 5; Length 568;
 Best Local Similarity 100.0%; Pred. No. 3.3e-52;
 Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKFCAGKTIREGMCQNSP 60
 DB 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKFCAGKTIREGMCQNSP 60

QY 61 DENTLDLVITNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEAL 120
 DB 61 DENTLDLVITNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEAL 120

QY 121 AGEQMIIITAGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVGTNATTITPGKWNLHR 180
 DB 121 AGEQMIIITAGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVGTNATTITPGKWNLHR 180

QY 181 MLRAAEYSVMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
 DB 181 MLRAAEYSVMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240

QY 241 DVQVCIIHTDVTNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTIMAGELNILPSSTT 300
 DB 241 DVQVCIIHTDVTNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTIMAGELNILPSSTT 300

QY 301 PTPIYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAMTSS 360
 DB 301 PTPIYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAMTSS 360

QY 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
 DB 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420

QY 421 YIGSVEEGKIADLVVWNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMFH 480
 DB 421 YIGSVEEGKIADLVVWNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMFH 480

QY 481 HGKAKFDTSIITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPK 540
 DB 481 HGKAKFDTSIITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPK 540

QY 541 TFEVFDGKLTCKTSPQVPLAQRYTFF 568
 DB 541 TFEVFDGKLTCKTSPQVPLAQRYTFF 568

RESULT 2

ADJ58251
 ID ADJ58251 standard; protein; 568 AA.
 XX
 AC ADJ58251;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE UreaseXY subunit #8.
 XX
 KW Immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
 XX
 OS Helicobacter felis.
 XX
 PN EP1176192-A2.
 XX
 PD 30-JAN-2002.
 XX
 PF 11-JUL-2001; 2001EP-00202666.
 XX
 PR 17-JUL-2000; 2000EP-00202565.
 XX
 PA (ALKU) AKZO NOBEL NV.
 XX
 PI Kusters JG, Cattoli G;
 XX
 DR WPI; 2002-124384/17.
 DR N-PSDB; ADJ58249.

XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
 PT the diagnosis of Helicobacter felis infections and in the preparation of
 PT vaccines.
 XX
 PS Disclosure; SEQ ID NO 15; 76pp; English.
 XX
 CC The present invention relates to a novel Helicobacter felis urease X and
 CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
 CC used in the manufacture of vaccines against Helicobacter felis infections
 CC and in diagnostic tests to detect antibodies against Helicobacter felis.
 CC Helicobacter felis is difficult to grow so it is more convenient to use
 CC the expression products of the genes encoding the urease X and Y subunits
 CC in the manufacture of vaccines. The present sequence represents an
 CC ureaseXY subunit of the invention.

XX Sequence 568 AA;

Query Match 99.1%; Score 2973; DB 5; Length 568;
 Best Local Similarity 98.9%; Pred. No. 1.1e-51;
 Matches 562; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKFCAGKTIREGMCQNSP 60
 DB 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKFCAGKTIREGMCQNSP 60

QY 61 DENTLDLVITNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEAL 120
 DB 61 DENTLDLVITNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEAL 120

QY 121 AGEQMIIITAGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVGTNATTITPGKWNLHR 180
 DB 121 AGEQMIIITAGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPDVGTNATTITPGKWNLHR 180

QY 181 MLRAAEYSVMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240
 DB 181 MLRAAEYSVMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADYEY 240

QY 241 DVQVCIIHTDVTNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTIMAGELNILPSSTT 300
 DB 241 DVQVCIIHTDVTNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTIMAGELNILPSSTT 300

QY 301 PTPIYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAMTSS 360
 DB 301 PTPIYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIIRPGSIAAEDVLHDMGVIAMTSS 360

QY 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
 DB 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420

QY 421 YIGSVEEGKIADLVVWNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMFH 480
 DB 421 YIGSVEEGKIADLVVWNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMFH 480

QY 481 HGKAKFDTSIITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPK 540
 DB 481 HGKAKFDTSIITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKFNDDTKAKITVDPK 540

QY 541 TFEVFDGKLTCKTSPQVPLAQRYTFF 568
 DB 541 TFEVFDGKLTCKTSPQVPLAQRYTFF 568

RESULT 3

ADJ58242
 ID ADJ58242 standard; protein; 568 AA.
 XX
 AC ADJ58242;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE UreaseXY subunit #2.
 XX

immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.

Helicobacter felis.

EP1176192-A2.

30-JAN-2002.

11-JUL-2001; 2001EP-00202666.

17-JUL-2000; 2000EP-00202565.

(ALKU) AKZO NOBEL NV.

Kusters JG, Cattoli G;

WPI; 2002-124384/17.

N-PSDB; ADJ58240.

Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.

Disclosure; SEQ ID NO 6; 76pp; English.

The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis. Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an ureaseXY subunit of the invention.

Sequence 568 AA;

Query Match 98.7%; Score 2960; DB 5; Length 568;

Best Local Similarity 98.6%; Pred. No. 2e-51;

Matches 560; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

1 MMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTTYGEELKFGAGKTIREGMQSNP 60

1 MMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTTYGEELKFGAGKTIREGMQSNP 60

61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120

61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120

121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTPGVDTNATTITPGKNLHR 180

121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTPGVDTNATTITPGKNLHR 180

181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

241 DVQVCIHTDTVNEAGYVDDTLNANGRAIHAYHLEGAGGHSPPVITMAGELNLPSTT 300

241 DVQVCIHTDTVNEAGYVDDTLNANGRAIHAYHLEGAGGHSPPVITMAGELNLPSTT 300

301 PTPYTTINTVAEHLDMLTCHLDRKIREDLQFSQRIIRPGSIAAEDVLHDMGVIA 360

301 PTPYTTINTVAEHLDMLTCHLDRKIREDLQFSQRIIRPGSIAAEDVLHDMGVIA 360

361 DSOAMGRAGEVIPRTWTADNKKKEFKLPEDGADNDNFRIKRYISKYTINPALTHGV 420

361 DSOAMGRAGEVIPRTWTADNKKKEFKLPEDGADNDNFRIKRYISKYTINPALTHGV 420

421 YIGSVEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOVYVREMG 480

421 YIGSVEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOVYVREMG 480

481 HGKAFDTSITFVSKVAYENGVEKGLERKVLVKNCRNITKKDFKFNKTAHITVDPK 540

Db 481 HGKAFDTSITFVSKVAYENGVEKGLERKVLVKNCRNITKKDFKFNKTAHITVDPK 540

QY 541 TFEVFDGKLCSTSKPTSQVPLAQRVTF 568

Db 541 TFEVFDGKLCSTSKPASEVPLAQRVTF 568

RESULT 4

ADJ58248

ID ADJ58248 standard; protein; 568 AA.

AC ADJ58248;

DT 06-MAY-2004 (first entry)

UreaseXY subunit #6.

immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.

OS Helicobacter felis.

PN EP1176192-A2.

PD 30-JAN-2002.

PF 11-JUL-2001; 2001EP-00202666.

PR 17-JUL-2000; 2000EP-00202565.

PA (ALKU) AKZO NOBEL NV.

PI Kusters JG, Cattoli G;

DR WPI; 2002-124384/17.

DR N-PSDB; ADJ58246.

Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.

Disclosure; SEQ ID NO 12; 76pp; English.

The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis. Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an ureaseXY subunit of the invention.

Sequence 568 AA;

Query Match 98.5%; Score 2953; DB 5; Length 568;

Best Local Similarity 98.2%; Pred. No. 2.8e-51;

Matches 558; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 MMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTTYGEELKFGAGKTIREGMQSNP 60

1 MMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTTYGEELKFGAGKTIREGMQSNP 60

61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120

61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120

121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTPGVDTNATTITPGKNLHR 180

121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTPGVDTNATTITPGKNLHR 180

181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

QY 241 DVQVCIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTWAGELNLPSTT 300
 DB 241 DVQVCIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTWAGELNLPSTT 300
 QY 301 PTIPYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
 DB 301 PTIPYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
 QY 361 DSQAMGRAGEVIPRTWTADNKKKEFGKLPEDGKNDNDFRIKRIYISKYTTINPALTGVSE 420
 DB 361 DSQAMGRAGEVIPRTWTADNKKKEFGKLPEDSADNDFRIKRIYISKYTTINPALTGVSE 420
 QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMMVFSEMGDSNASVPTPQPVYVREMFH 480
 DB 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMMVFSEMGDSNASVPTPQPVYVREMFH 480
 QY 481 HGKAKFDTSTIFVSKVAYENGVEKLGLEKROVLVKNCRNITKKDFKENDTAKITVDPK 540
 DB 481 HGKAKFDTSTIFVSKVAYENGVEKLGLEKRLKVLVKNCRNITKKDFKFNKTAHITVDPK 540
 QY 541 TFEVFDGKLCSTKSPASEVPLAQRYPFF 568
 DB 541 TFEVFDGKLCSTKSPASEVPLAQRYPFF 568

RESULT 5
 ADJ58245
 ID ADJ58245 standard; protein; 496 AA.
 AC ADJ58245;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE UreaseX subunit #4.
 KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
 XX
 OS Helicobacter felis.
 XX
 PN EPI176192-A2.
 XX
 PD 30-JAN-2002.
 XX
 PF 11-JUL-2001; 2001EP-00202666.
 XX
 PR 17-JUL-2000; 2000EP-00202565.
 XX
 PA (ALKU) AKZO NOBEL NV.
 XX
 PI Kusters JG, Cattoli G;
 XX
 DR WPI; 2002-124384/17.
 DR N-PSDB; ADJ58243.
 XX
 PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
 the diagnosis of Helicobacter felis infections and in the preparation of
 vaccines.
 XX
 PS Disclosure; SEQ ID NO 9; 76pp; English.
 XX
 CC The present invention relates to a novel Helicobacter felis urease X and
 Y subunit polypeptides and immunogenic fragments. The polypeptides are
 used in the manufacture of vaccines against Helicobacter felis infections
 and in diagnostic tests to detect antibodies against Helicobacter felis.
 CC Helicobacter felis is difficult to grow so it is more convenient to use
 the expression products of the genes encoding the urease X and Y subunits
 in the manufacture of vaccines. The present sequence represents an
 ureaseX subunit of the invention.
 CC
 SQ Sequence 496 AA;

Query Match 86.4%; Score 2591; DB 5; Length 496;

Best Local Similarity 99.2%; Pred. No. 4.3e-44;
 Matches 491; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 1 MKMKQEVYNTYGTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMGQSNP 60
 DB 1 MKMKQEVYNTYGTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMGQSNP 60
 QY 61 DENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNQMKGVSHPMVVGVGTEAL 120
 DB 61 DENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNQMKGVSHPMVVGVGTEAL 120
 QY 121 AGEGLIITAGGIDSHTHFLSPQOFPALTANGVTTMPFGGTPVDGNTATTITPGKNLHR 180
 DB 121 AGEGLIITAGGIDSHTHFLSPQOFPALTANGVTTMPFGGTPVDGNTATTITPGKNLHR 180
 QY 181 MLRAAEYSMNVEFGKGNSSSKQLVEQVEAGAIGFKLHEDWGTTPPSAIDHCLSVADY 240
 DB 181 MLRAAEYSMNVEFGKGNSSSKQLVEQVEAGAIGFKLHEDWGTTPPSAIDHCLSVADY 240
 QY 241 DVQVCIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTWAGELNLPSTT 300
 DB 241 DVQVCIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPDVTWAGELNLPSTT 300
 QY 301 PTIPYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
 DB 301 PTIPYTTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
 QY 361 DSQAMGRAGEVIPRTWTADNKKKEFGKLPEDGKNDNDFRIKRIYISKYTTINPALTGVSE 420
 DB 361 DSQAMGRAGEVIPRTWTADNKKKEFGKLPEDGKNDNDFRIKRIYISKYTTINPALTGVSE 420
 QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMMVFSEMGDSNASVPTPQPVYVREMFH 480
 DB 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMMVFSEMGDSNASVPTPQPVYVREMFH 480
 QY 481 HGKAKFDTSTIFVSKVAYENGVEKLGLEKROVLVKNCRNITKKDFKENDTAKITVDPK 540
 DB 481 HGKAKFDTSTIFVSKVAYENGVEKLGLEKRLKVLVKNCRNITKKDFKFNKTAHITVDPK 540
 RESULT 6
 ADQ37851
 ID ADQ37851 standard; protein; 569 AA.
 XX
 AC ADQ37851;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE H. bizzozeronii ureB polypeptide.
 XX
 KW Urease; urease gene cluster; urease structural gene;
 KW urease accessory gene; ureB; Helicobacter bizzozeronii infection;
 KW antibacterial; enzyme.
 XX
 OS Helicobacter bizzozeronii.
 XX
 PN US2004142343-A1.
 XX
 PD 22-JUL-2004.
 XX
 PF 12-AUG-2003; 2003US-00639273.
 XX
 PR 16-AUG-2002; 2002US-0404337P.
 XX
 PA (CHAN/) CHANG Y.
 PA (SIMP/) SIMPSON K W.
 PA (ZHUJ/) ZHU J.
 XX
 PI Chang Y, Simpson KW, Zhu J;
 XX
 DR WPI; 2004-533502/51.
 DR N-PSDB; ADQ37850.
 XX

PT Novel isolated nucleic acid molecule having urease gene cluster, and
PT conferring on Helicobacter bizzozeronii ability to produce urease, useful
PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
XX
XX PS
XX Claim 24; SEQ ID NO 5; 40pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule conferring on
CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
CC acid molecule is a urease gene cluster comprising at least one urease
CC structural gene and at least one urease accessory gene. The nucleic acid
CC molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
CC invention also relates to an isolated protein encoded by the nucleic
CC acid, a vaccine for preventing onset of disease in mammals infected by H.
CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
CC antibody or its binding portion raised against the nucleic acid. The
CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
CC against onset of disease caused by infection of H. bizzozeronii, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizzozeronii in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridization assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizzozeronii is
CC present in the sample. This sequence represents the H. bizzozeronii ureB
CC polypeptide.
XX
XX SQ Sequence 569 AA;

Query Match 77.6%; Score 2326; DB 8; Length 569;
Best Local Similarity 74.1%; Pred. No. 1.2e-38;
Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;
QY 2 KMKQEVNTYPTKGDVRLGDTDLAEVHEHDYTYGEEELKFGAGKTIREGMSQNSPD 61
DB 3 KISRKEYVMYPTTGDVRLGDTDLAEVHEHDYTYGEEELKFGAGKTIREGMSQNSPD 62
QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMMDQGVSPHMVGVGTEALA 121
DB 63 SHELDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMMDQGVSPHMVGVGTEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVYTMFGGTTGPDGTNATITPGKNNLHRM 181
DB 123 AEGLIIVTAGGIDSHTHFLSPQFPPTALANGVYTMFGGTTGPDGTNATITPGKNNLHRM 182
QY 182 LRAAEYSMNVGFLGKGNSSKKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 241
DB 183 LRAAEYSMNVGFLGKGNSSKKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 242
QY 242 VQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTWAGELNIPSTTP 301
DB 243 VQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTWAGELNIPSTTP 302
QY 302 TIPTINTVAEHLDMVCHLDRNIREDLQFSQSRIRPGSIAEDVLHDMGVIAMTSSD 361
DB 303 TIPTINTVAEHLDMVCHLDRNIREDLQFSQSRIRPGSIAEDVLHDMGVIAMTSSD 362
QY 362 SOAMGRAGEVPIRTWQADKNKKEFGKLPEDKNDNFRIKRYISKYITINPAIHGISEY 421
DB 363 SOAMGRAGEVPIRTWQADKNKKEFGKLPEDKNDNFRIKRYISKYITINPAIHGISEY 422
QY 422 IGSVEEGKIADLVNPAFFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFGHH 481
DB 423 IGSVEEGKIADLVNPAFFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFGHH 482
QY 482 GKAKFDTSIITFVSKVAYENGVEKELQVLPVKNCRNITKKDFKNDKTAKITVDPKT 541
DB 483 GKAKFDTSIITFVSKVAYENGVEKELQVLPVKNCRNITKKDFKNDKTAKITVDPKT 542
QY 542 FEYFVDSKICTSKPTSQVPLAORTFF 568
DB 543 YKVKVDGKEVTSKAADKISLAQLNLF 569

RESULT 7
AAR04580
ID AAR04580 standard; protein; 569 AA.
XX
XX AC AAR04580;
XX
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 21-SEP-1990 (first entry)
XX Part of protein with urease activity.
XX Urease; probe.
XX Helicobacter pylori.
XX WO9004030-A.
XX 19-APR-1990.
XX 06-OCT-1988; 88FR-00013135.
XX 06-OCT-1988; 88FR-00013135.
XX (INSP) INST PASTEUR.
XX (INRM) INSERM INST NAT SANTE & RECH MED.
XX Labigne A;
XX WPI: 1990-147844/19.
XX N-PSDB; AAQ04329.
XX New nucleotide sequences encoding Campylobacter pylori-urease - and
XX derived vectors, transformants, protein, antibodies and probes, useful in
XX diagnosis, treatment and prevention of infections.
XX Claim 12; Page 34; 47pp; French.
XX The protein can be used for the prodn. of antibodies and to prepare
XX vaccines for the prevention /treatment of C. pylori infections. See also
XX AAR04578-81. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24
XX -OCT-2003 to standardise OS field)
XX SQ Sequence 569 AA;
Query Match 76.6%; Score 2297; DB 2; Length 569;
Best Local Similarity 73.4%; Pred. No. 4.5e-38;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKQEVNTYPTKGDVRLGDTDLAEVHEHDYTYGEEELKFGAGKTIREGMSQNSPD 61
DB 3 KISRKEYVMYPTTGDVRLGDTDLAEVHEHDYTYGEEELKFGAGKTIREGMSQNSPD 62
QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMMDQGVSPHMVGVGTEALA 121
DB 63 KEELDLITNALIVDYTYGIYKADIGIKNGKIHGKAGNKMMDQGVSPHMVGVGTEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVYTMFGGTTGPDGTNATITPGKNNLHRM 181
DB 123 GEGMIITAGGIDSHTHFLSPQFPPTALANGVYTMFGGTTGPDGTNATITPGKNNLHRM 182
QY 182 LRAAEYSMNVGFLGKGNSSKKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 241
DB 183 LRAAEYSMNVGFLGKGNSSKKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 242
QY 242 VQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTWAGELNIPSTTP 301
DB 243 VQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTWAGELNIPSTTP 302
QY 302 TIPTINTVAEHLDMVCHLDRNIREDLQFSQSRIRPGSIAEDVLHDMGVIAMTSSD 361
DB 303 TIPTINTVAEHLDMVCHLDRNIREDLQFSQSRIRPGSIAEDVLHDMGVIAMTSSD 362

QY 362 SOAMGRAGEVTPRTWQTADKXKKEFGKLPEDGKNDNFRIKRYISKYTYINPALTHGVSEY 421
 DB 363 SOAMGRVGEVITRTWQTADKXKKEFGKLPEDGKNDNFRIKRYISKYTYINPAIAHGISY 422
 QY 422 IGSVEEGKIADLVVWNPAPFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYTYREMGHH 481
 DB 423 VGSVEGKVADLVWSPAPFFGVKPNMIKGGFIALSQMGDANASIPTPQPVYTYREMFHH 482
 QY 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLPVKNCRNITKDFKNDKTAKITVDPKT 541
 DB 483 GKAKYDANITFVSOAAYDKGKEELGLERQVLPVKNCRNITKDMQFNTTAHIEVNPET 542
 QY 542 FEVFDGKLCSTKSPQVPLAQRITFF 568
 DB 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569

RESULT 8
 AAW07194
 ID AAW07194 standard; protein; 569 AA.
 AC AAW07194;
 XX AC
 XX AC
 DT 16-OCT-2003 (revised)
 DT 11-FEB-1997 (first entry)
 XX DT
 DE H. pylori urease B subunit.
 XX DE
 XX DE
 KW Urease; ureA gene; ureB gene; vaccine.
 XX KW
 XX KW
 OS Helicobacter pylori; strain CPM630.
 XX OS
 XX OS
 PN W09633732-A1.
 XX PN
 PD 31-OCT-1996.
 XX PD
 PF 25-APR-1996; 96WO-US005800.
 XX PF
 XX PF
 PR 28-APR-1995; 95US-00431041.
 PR 06-DEC-1995; 95US-00568122.
 XX PR
 XX PR
 PA (ORAV-) ORAVAX INC.
 XX PA
 XX PA
 PI Lee CK, Monath TP, Ackerman SK, Thomas WD, Kleanthous H;
 PI Weltzin RA, Pappo J, Ermak T, Guirakhoo F, Bhagat H, Sussman I;
 XX PI
 XX PI
 DR WPI; 1996-497373/49.
 DR N-PSDB; AAT44351.
 XX DR
 XX DR
 PT Vaccine for inducing mucosal response to Helicobacter - contg. multimeric
 PT urease complex and pref. an antibiotic, anti-secretory agent or bismuth
 salt.
 XX PT
 XX PT
 PS Disclosure; Page 71-73; 98pp; English.
 XX PS
 XX PS
 CC Urease A (AAW07193) and B (AAW07194) subunits of Helicobacter pylori are
 CC encoded by the ureA + ureB gene locus of clinical isolate CPM630.
 CC CC
 CC Vectors, e.g. pORV214 (see also AAT44351), carrying the gene locus can be
 CC used to transform host cells for the large-scale prodn. of recombinant,
 CC enzymatically inactive, multimeric urease. The urease complex induces a
 CC mucosal immune response that can treat or prevent Helicobacter, esp. H.
 CC pylori, gastroduodenal infection. (Updated on 16-OCT-2003 to standardise
 CC OS field)
 XX CC
 XX CC
 SQ Sequence 569 AA;
 Query Match 76.6%; Score 2297; DB 2; Length 569;
 Best Local Similarity 73.4%; Pred. No. 4.5e-38;
 Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMVKQEVYNTYPTKDKVRLGDTDLAEVHEHDYTYTGEEELKFGAGKTYIREMGQNSPD 61
 DB 3 KISRKEYVSMYPTTGDKVRLGDTDLAEVHEHDYTYTGEEELKFGAGKTYIREMGQNSPD 62

QY 62 ENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDQGVSPHVVGVGTEALA 121
 DB 63 KEELDLIITNALIVDTGIYKADIGIKDGKAGIKGGNKMDQGVKNNLSVGPATEALA 122
 QY 122 GEGMIITAGGIDSHFSLSPQFPPTALANGVTTMFGGGTGPDVGTNATITTPGKNLHRM 181
 DB 123 GEGLVITAGGIDTHFISPOQIPTAFASGVTTMTIGGGTGPADGTNATITTPGRRNLKWM 182
 QY 182 LRAAEYSMNUGFLKGNSSSKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEYD 241
 DB 183 LRAAEYSMNUGFLKGNASNDASLADQIEAGAGFKIHEDWGTTPSAINHADVADKYD 242
 QY 242 VQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPVITMAGELNLPSTTP 301
 DB 243 VQVAIHTDTLNEAGCVEDTMAAIAGRMTHTFHTGAGGCGHAPDIIKVAGEHNILPASTNP 302
 QY 302 TIPTTINTVAEHLDMLTCHLDKIREDLQFSQSRIRPGSIAAEVDLHDMGVIAMTSSD 361
 DB 303 TIPTTINTVAEHLDMLTCHLDKIREDLQFSQSRIRPGSIAAEVDLHDMGVIAMTSSD 362
 QY 362 SOAMGRAGEVTPRTWQTADKXKKEFGKLPEDGKNDNFRIKRYISKYTYINPALTHGVSEY 421
 DB 363 SOAMGRVGEVITRTWQTADKXKKEFGKLPEDGKNDNFRIKRYISKYTYINPAIAHGISY 422
 QY 422 IGSVEEGKIADLVVWNPAPFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYTYREMGHH 481
 DB 423 VGSVEGKVADLVWSPAPFFGVKPNMIKGGFIALSQMGDANASIPTPQPVYTYREMFHH 482
 QY 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLPVKNCRNITKDFKNDKTAKITVDPKT 541
 DB 483 GKAKYDANITFVSOAAYDKGKEELGLERQVLPVKNCRNITKDMQFNTTAHIEVNPET 542
 QY 542 FEVFDGKLCSTKSPQVPLAQRITFF 568
 DB 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569

RESULT 9
 AAW98511
 ID AAW98511 standard; protein; 569 AA.
 AC AAW98511;
 XX AC
 XX AC
 DT 31-MAR-1999 (first entry)
 XX DT
 XX DT
 DE H. pylori GHPO 1248 protein.
 XX DE
 XX DE
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 KW peptic ulcer disease.
 XX KW
 XX KW
 OS Helicobacter pylori.
 XX OS
 XX OS
 PN W09843478-A1.
 XX PN
 XX PN
 PD 08-OCT-1998.
 XX PD
 PF 01-APR-1998; 98WO-US006371.
 XX PF
 XX PF
 PR 01-APR-1997; 97US-00833457.
 PR 24-JUN-1997; 97US-00881227.
 PR 29-JUL-1997; 97US-00902615.
 XX PR
 XX PR
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX PA
 XX PA
 PI Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
 XX PI
 XX PI
 DR WPI; 1998-542293/46.
 DR N-PSDB; AAX14230.
 XX DR
 XX DR
 PT New isolated Helicobacter polynucleotides - used to develop products for
 the diagnosis, prevention and treatment of Helicobacter infections and

PT		gastrointestinal diseases.
XX		Claim 8; Page 960-962; 2054pp; English.
PS		This sequence represents a Helicobacter pylori GHPD protein of the
XX		invention. The polypeptides can be used for preventing or treating
CC		Helicobacter infections, and gastroduodenal diseases associated with
CC		these infections, including acute, chronic, and atrophic gastritis, and
CC		peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC		used for the production of antibodies. The products can also be used for
CC		detection and diagnosis
XX		Sequence 569 AA;
SQ		
	Query Match	76.6%; Score 2297; DB 2; Length 569;
	Best Local Similarity	73.4%; Pred. No. 4.5e-38;
	Matches 416; Conservative	76; Mismatches 75; Indels 0; Gaps 0;
Qy	2 KMKKQEVNTYGPFGDKVRLGDTDLMAEVEHDVTYYTGELKFAGKTIREGMQSQNSPD	61
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	3 KISRKEYVSVMYGPTTGDVKVRLGDTDLAEVEHDVTYYTGELKFAGKTIREGMQSQNSPD	62
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	62 ENTLDLVITWAMIIDYGIYKADIGIKNGKHGIGKAGNKMDQGVSPHVVVGVTALTA	121
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	63 KBEELDIITNALIVDYTYGIYKADIGIKDGKTAGIKGGNKMDQGVXNNLSVGPATATA	122
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	122 GEGMIITPAGGISHTFELSPOOPFTALANGVYTMFGCGTGVDCTNATTTPGKNLHRM	181
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	123 GEGLIIVTAGIDTHIHIFISQQIPTAFASGVTTMGGTGPGADGNATTTPGRRNLKWM	182
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	182 LRAAEEYSNMNVFLGKGNSSKKQLVQEVEAGATGFKLHEDWGTPPSAIDHCLSVADEYD	241
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	183 LRAAEEYSNMNLGFLAKGNASNDSLASDAQIEAGAIGFKIHEHDWGTTGPSAINHALDVADKYD	242
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	242 VQVCIIHTDTVNEAGYDDTLNAMNGRAIHAYHIEGACGHSPDVI TWAGELNILPSSTTP	301
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	243 VOVAIHDTTLNEAGCVEDTWAAIAGRTMTHTFHTEGAGGSHAPDI IKVAGEHNILPASTNP	302
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	302 TIPVTINTVAEHLDMLTCCHLDKR IREDLFQSOSRIRPGSIAAEDVLHDMGVIAMTSSD	361
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	303 TIFPTVNTAEHMDMLVCHLDKSIKEDEVQFADSRIRPOTIAAEDTLHDMGFISTSSD	362
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	362 SQAMGRAGEVIPRTWQTADNKKKEFGKLPGEDGKDNDNFRIKRYISKYTINPALTHGVSEY	421
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	363 SQAMGRAGEVITRTWQTADNKKKEFGKLKEEKGDNDNFRIKRYLSKYTINPAIAHGISEY	422
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	422 IGVSVEGKIADLVVNNDPAFPGVZPKIVIKGMVVFSEMGSDSNASVPPTQPVTREMPGHH	481
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	423 VGSVEVGKVADLVMSPAFFGVKPNMIIKGFIALSQMGDANASIPTQPVTREMPFAHH	482
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	482 GKAKFTDTSITFVSKVAENVGKKGLGLEROVL PVKNCRNITKKDFKENDKTKATITVDPKT	541
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	483 GKAKYDANITTFVSOAAVDKGIKBELGLERQVL PVKNCRNITKKDMQNFDTTAHLEVPET	542
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	542 FEYFVDGKLTCKTSKPTSQVPLAQRYTF 568	
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
Qy	543 YHFVVDGKVTSPANKVSLAQLESIF 569	
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
RESULT 10		
ABUJ30687		
ID	ABUJ30687 standard; protein; 569 AA.	
XX		
AC	ABUJ30687;	
XX		
DT	19-JUN-2003 (first entry)	
XX		
DE	Protein encoded by Prokaryotic essential gene #16214.	
XX		
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.	
OS		
XX	Helicobacter pylori.	
XX		

WO200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WO-US09107.
21-MAR-2001; 2001US-0081542.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.
N-PSDB; ACA34557.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 58611; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at fp.wipo.int/pub/published_pct_sequences

Sequence 569 AA;

Query Match 76.6%; Score 2297; DB 6; Length 569;
Best Local Similarity 73.4%; Pred. No. 4.5e-38;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0

Qy 2 KMKKQEYVNTYGPTGDKVKRLGDFTDLMAEVEHDYTTTGEELKFGAGKTIREGMQSNSPD 61
| : : |
Db 3 KISRKEYSVSMYGP TTGDKVKRLGDFTDLAEVEHDTTYTTEELKFGGGKLTREGMSQSNPS 62
| : : |
Qy 62 ENTLDLIVTNALIIDYGIKYADIGIKNGKTHGKGANKDMQDGVSPHMVVGVGTALA 121
| : : |
Db 63 KEEDLLIITNALIVDYTIYGIKYADIGIKDGKTAGIKGNKMOMQDGVKNVL SVGPATEALA 122
| : : |
Qy 122 GEGMIITAGIDISHTFPLSQPF TALANGVTMTMFGGSGTGVPDGTNATTTTTPGKWNLRH 181
| : : |

Db 123 GEGLVITAGGIDTHIHFIPOQITAFASGVTTMIGGTGPDGNTATTTPGRRLKWM 182
 QY 182 LRAAEYSNMVGFLLKGNSSKKQLVEQVAGAGIAGFKLHEDWGTTPSAIDHCLSVADYD 241
 Db 183 LRAAEYSNMVGFLLKGNSSKKQLVEQVAGAGIAGFKLHEDWGTTPSAIDHCLSVADYD 242
 QY 242 VQVCIHTDTVNEAGYVDDTLNMGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 301
 Db 243 VQVCIHTDTVNEAGYVDDTLNMGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 302
 QY 302 TIPVTINTVAEHLDMVCHHLDKRIEDLOFSQSRIRPGSIAAEDVLDHMGVIAWTSDD 361
 Db 303 TIPVTINTVAEHLDMVCHHLDKRIEDLOFSQSRIRPGSIAAEDVLDHMGVIAWTSDD 362
 QY 362 SQAMGRAGEVITPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
 Db 363 SQAMGRAGEVITPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 422
 QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFHH 481
 Db 423 VGSVEGKVADLVWSPAFFGKPNMIKGGFIALSQMGDANASIPTPQPVYVREMFHH 482
 QY 482 GKAKFDTISITFVSKVAYENGVEKELGLERQVLPVKNCNITKKDPKNDKTAKITVDPKT 541
 Db 483 GKAKFDTISITFVSKVAYENGVEKELGLERQVLPVKNCNITKKDPKNDKTAKITVDPKT 542
 QY 542 FEVFDGKCLTSKPTSOVPLAQRVTF 568
 Db 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569

RESULT 11
 ID ADU05536 standard; protein; 569 AA.
 AC ADU05536;
 DT 27-JAN-2005 (first entry)
 XX H. pylori antigenic protein HP0072.
 DE antibacterial; antigenic; H. pylori infection.
 KW Helicobacter pylori.
 OS Location/Qualifiers
 PH 533..567
 FT Region
 FT /note= "Immunogenic region. Region specifically claimed
 in claim 14"
 XX WO2004094467-A2.
 XX 04-NOV-2004.
 XX 22-APR-2004; 2004WO-EP004255.
 XX 22-APR-2003; 2003EP-00450097.
 XX (INTE-) INTERCELL AG.
 XX Meinke A, Min Bui D, Nagy E, Henics T;
 DR WPI; 2004-775908/76.
 DR N-PSDB; ADU05358.
 XX New hyperimmune serum reactive antigens from Helicobacter pylori, and
 PT encoding nucleic acid molecules, useful for diagnosing, preventing or
 PT treating H. pylori infections.
 XX Claim 14; SEQ ID NO 184; 176pp; English.
 PS The invention relates to an isolated nucleic acid molecule encoding a

hyperimmune serum reactive antigen or its fragment. The composition
 (including the nucleic acid molecule, hyperimmune serum-reactive antigen
 or antibody) is useful for manufacturing a medicament or pharmaceutical
 preparations (e.g. a vaccine) for treating or preventing H. pylori
 infections. The antigen or its fragment may also be used for isolating,
 purifying and/or identifying an interaction partner of the hyperimmune
 serum reactive antigen or fragment; for generating a peptide binding to
 the hyperimmune serum reactive antigen or fragment, where the peptide is
 selected from aptamers and spiegelmers; or for manufacturing a functional
 selected from aptamers and spiegelmers; or for manufacturing a functional
 ribonucleic acid selected from ribozymes, antisense nucleic acids and
 siRNA. The present sequence represents the amino acid sequence of an H.
 pylori antigenic protein.
 CC Sequence 569 AA;
 SQ
 Query Match 76.6%; Score 2297; DB 8; Length 569;
 Best Local Similarity 73.4%; Pred. No. 4.5e-38;
 Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
 QY 2 KMKQEQYNTYGTPTKDGKVRGLGDTDLMAVEHDYTYGEEELKFGAGKTIREGMSQNSPD 61
 Db 3 KISRKEYVSMYGTPTGDKVRGLGDTDLMAVEHDYTYGEEELKFGAGKTIREGMSQNSPD 62
 QY 62 ENTLDLVTNAMIIDYTGIIYKADIGIKNGKIHGIGKAGKQMDGVSPHMVVGVTGTEALA 121
 Db 63 KEELDLIITNALIVDYGIIYKADIGIKNGKIHGIGKAGKQMDGVSPHMVVGVTGTEALA 122
 QY 122 GEGMLITAGGIDSHTHFLSPQFPFALANGVTTMFGGGTGPVDGNTATTTPCKKNLHRM 181
 Db 123 GEGMLITAGGIDSHTHFLSPQFPFALANGVTTMFGGGTGPVDGNTATTTPCKKNLHRM 182
 QY 182 LRAAEYSNMVGFLLKGNSSKKQLVEQVAGAGIAGFKLHEDWGTTPSAIDHCLSVADYD 241
 Db 183 LRAAEYSNMVGFLLKGNSSKKQLVEQVAGAGIAGFKLHEDWGTTPSAIDHCLSVADYD 242
 QY 242 VQVCIHTDTVNEAGYVDDTLNMGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 301
 Db 243 VQVCIHTDTVNEAGYVDDTLNMGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 302
 QY 302 TIPVTINTVAEHLDMVCHHLDKRIEDLOFSQSRIRPGSIAAEDVLDHMGVIAWTSDD 361
 Db 303 TIPVTINTVAEHLDMVCHHLDKRIEDLOFSQSRIRPGSIAAEDVLDHMGVIAWTSDD 362
 QY 362 SQAMGRAGEVITPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
 Db 363 SQAMGRAGEVITPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 422
 QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFHH 481
 Db 423 VGSVEGKVADLVWSPAFFGKPNMIKGGFIALSQMGDANASIPTPQPVYVREMFHH 482
 QY 482 GKAKFDTISITFVSKVAYENGVEKELGLERQVLPVKNCNITKKDPKNDKTAKITVDPKT 541
 Db 483 GKAKFDTISITFVSKVAYENGVEKELGLERQVLPVKNCNITKKDPKNDKTAKITVDPKT 542
 QY 542 FEVFDGKCLTSKPTSOVPLAQRVTF 568
 Db 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569
 RESULT 12
 ID ADS09177 standard; protein; 568 AA.
 AC ADS09177;
 DT 16-DEC-2004 (first entry)
 XX H. pylori urease beta subunit.
 DE Urease; beta subunit; UreB; gastric ulcer; peptic ulcer; gastritis;
 KW stomach cancer; vaccine; antibody; immune reaction.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:43:17 ; Search time 51 Seconds
(without alignments)
974.851 Million cell updates/sec

Title: US-09-904-994b-3

Perfect score: 2999

Sequence: 1 MKMKQBYVNTYGTGDKV.....KLCTSKPTSQVPLAQRVTF 568

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC Cellerai SID33/prodata/2/iaa/5_COMB.pap:*
- 2: /EMC Cellerai SID33/prodata/2/iaa/6_COMB.pap:*
- 3: /EMC Cellerai SID33/prodata/2/iaa/7_COMB.pap:*
- 4: /EMC Cellerai SID33/prodata/2/iaa/H_COMB.pap:*
- 5: /EMC Cellerai SID33/prodata/2/iaa/PTUS_COMB.pap:*
- 6: /EMC Cellerai SID33/prodata/2/iaa/RE_COMB.pap:*
- 7: /EMC Cellerai SID33/prodata/2/iaa/backfiles.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2304	76.8	568	2	US-09-338-920B-12
2	2297	76.6	569	2	US-09-338-920B-6
3	2294	76.5	569	1	US-08-467-822-26
4	2294	76.5	569	2	US-08-432-697-26
5	2294	76.5	569	2	US-08-466-248-26
6	2290	76.4	566	1	US-08-920-095-3
7	2290	76.4	566	5	PCT-US96-05800-3
8	2282	76.1	570	2	US-09-431-705-5
9	2244	74.8	569	1	US-08-467-822-21
10	2244	74.8	569	2	US-08-432-697-21
11	2244	74.8	569	2	US-08-466-248-21
12	2244	74.8	569	2	US-09-338-920B-10
13	2237	74.6	559	2	US-09-338-920B-8
14	2044	68.2	568	1	US-07-732-242C-3
15	1834	61.2	566	2	US-09-252-991A-26887
16	1831	61.1	569	1	US-08-467-822-27
17	1831	61.1	569	2	US-08-432-697-27
18	1831	61.1	569	2	US-08-466-248-27
19	1831	61.1	569	2	US-09-543-681A-6029
20	1830	61.0	625	2	US-09-489-039A-9216
21	1829	61.0	567	2	US-09-338-352-5912
22	1826	60.9	567	1	US-08-967-513-5
23	1826	60.9	567	1	US-08-687-645B-5
24	1804	60.2	571	2	US-09-710-279-118
25	1804	60.2	840	1	US-08-467-822-25
26	1804	60.2	840	2	US-08-432-697-25

27	1804	60.2	840	2	US-08-466-248-25	Sequence 25, Appl
28	1803	60.1	573	2	US-09-134-001C-5026	Sequence 5026, Ap
29	1720	57.4	570	2	US-09-602-777A-14	Sequence 14, Appl
30	1631	54.4	534	2	US-09-602-777A-16	Sequence 16, Appl
31	1166	38.9	308	2	US-09-338-920B-14	Sequence 14, Appl
32	469	15.6	121	2	US-10-012-819-160	Sequence 160, Appl
33	345	11.5	2142	2	US-09-540-236-3459	Sequence 3459, Ap
34	333	11.1	1078	2	US-09-583-110-4036	Sequence 4036, Ap
35	333	11.1	1080	2	US-09-107-433-4843	Sequence 4843, Ap
36	331	11.0	956	2	US-09-134-078-63	Sequence 63, Appl
37	328	10.9	1188	2	US-09-206-942-63	Sequence 59, Appl
38	328	10.9	1188	2	US-10-193-764-59	Sequence 65, Appl
39	327	10.9	1180	2	US-09-206-942-65	Sequence 61, Appl
40	327	10.9	1180	2	US-10-193-764-61	Sequence 10606, A
41	324	10.8	622	2	US-09-949-016-10606	Sequence 69, Appl
42	321	10.7	1095	2	US-09-206-942-69	Sequence 65, Appl
43	321	10.7	1095	2	US-10-193-764-65	Sequence 2, Appli
44	321	10.7	1536	1	US-08-038-682-2	Sequence 2, Appli
45	321	10.7	1536	1	US-08-302-832-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-338-920B-12
; Sequence 12, Application US/09338920B
; Patent No. 6709851

; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.

; TITLE OF INVENTION: Stabilization of Helicobacter Urease

; FILE REFERENCE: 06132/023002

; CURRENT APPLICATION NUMBER: US/09/338,920B

; CURRENT FILING DATE: 2002-12-23

; PRIOR APPLICATION NUMBER: US 08/928,081

; PRIOR FILING DATE: 1997-09-12

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 568

; TYPE: PRT

; ORGANISM: Helicobacter heilmannii

US-09-338-920B-12

Query Match	76.8%	Score 2304;	DB 2;	Length 568;
Best Local Similarity	73.7%;	Pred. No. 1e-40;		
Matches 418;	Conservative 74;	Mismatches 74;	Indels 1;	Gaps 1;
QY	2	MMKKQBYVNTYGTGDKVRLGDTDLWAEVHDYTYGELKFGAGKTIREGMGQNSPD	61	
Db	3	KISRKEYVSMYGTGDKVRLGDTDLLEVEHDTTYGEEIKFGGKTIRDGMGQNSPS	62	
QY	62	ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVEALA	121	
Db	63	SHELDLVTNALTIVDTYGIYKADIGIKNGKIHGKAGNKLDQGVNRLCVGPATEALA	122	
QY	122	GEGLMIITAGGIDSHYFLSPQFPPTALANGVTMTGGTGPVDGTTNATITPGKWLHRM	181	
Db	123	AEGLVITAGGIDTHIFHSPOQIPTAFASGITMTGGTGPADGTTNATITPGRWLKEK	182	
QY	182	LRAAEYSNVGFLGKGNSSKKQLVEQVAGAGIFKLHEDMGTTPTSAIDHCLSVADEYD	241	
Db	183	LRASEYAMNUGLYLGNVSPFALIDQEAAGIFKIHEDMGSTPSAINHAIADKYD	242	
QY	242	VQVCITHTVNEAGYVDDTLANMNGRAIHAYHIEGAGGSHSPDVTITMAGELNIPSTTP	301	
Db	243	VQVALHTDTLNEAGCVEDTLEAIGRTIHTFTEGAGGGHAPDVIKWAGEFNILPASTNP	302	
QY	302	TIPYTIINTVAEHLDMTCHHLDKRIREDLQFSQSRIRPGSTAAEDVLHDMGVIAWTSSD	361	
Db	303	TIPFTKNTAEAHMDMLM-CHHLDKNIKEDVEFADSRIRPQTIAAEDKLHDMGIFITSDD	361	

QY 362 SOAMGRAGEV1PRTWQTADKNKKEFGKLPEDGKONDNPRIKRYISKYKTINPALTHGVSEY 421
Db 362 SOAMGRVGEV1TRTWQTADKNKKEFGRLPEEKGDNDNFR1KRYISKYKTINPAITHGISEY 421
QY 422 IGSVEEGKIADLVVNPAPFGVKPKIV1KGGMVVFSMGDSNASVPTPOPVYVYREMFHGH 481
Db 422 VGSVEGKIADLVVNSPAFFGKPNMI1KGGFIALSQMGDANAS1PTPOPVYVYREMFHGH 481
QY 482 GKAKFDTSITFVSKVAYENGVEKLGRLERQVLVPVKNCRNITKQDFKNDKTAKITVDPKT 541
Db 482 GKAKFDNITFVSVQVAYENG1KHELGLQRLVLPVKNCRNITKQDLKFNVDVTAHIEVNPET 541
QY 542 FEVVDGKLCSTKPTSQVPLAQRVTF 569
Db 542 YKVKVDGNEVTSAAADKLSLAQLNLF 568
RESULT 2
US-09-338-920B-6
; Sequence 6, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-338-920B-6
Query Match 76.6%; Score 2297; DB 2; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.4e-40;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKKQEVNTYPTKGDVRLGDTDLMAVEVDHYTTYGEELKFGAGKTIREGMSQNSPD 61
Db 3 KISRKEVSMYGPPTGDKVRLGDTDLIAVEVDHYTYGEELKFGGKTIREGMSQNSPS 62
QY 62 ENTLDLVITNAMIIDYGIYKADIGIKNGKIHGKAGKNQMDGVSPHVVGVGTAL 121
Db 63 KEELDLIITNALIVDYTYGIYKADIGIKGKTIAGIGKGNQMDGVKNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQPTALANGVTMTFGGTGPVDTGNATITPGKNLHRM 181
Db 123 GEGLVITAGGIDTHIFHSPOQIPTAFASGVYTWIGGTGPADGTNATITPGRENLRKM 182
QY 182 LRAAEYSVMNVGFLGKNSKKOLVEVEAGAIGFKLHEDWGTTPSAIDHCLSVADVD 241
Db 183 LRAAEYSVMNVGFLGKNSNASLADQIEAGAIKFKIHEDWGTTPSAINHALDVADKYD 242
QY 242 VQVCITDVTNEAGVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNLPSSPTP 301
Db 243 VQVAITDVTNEAGCVEDTMAAIGRTMHTFTEGAGGHPADIIKVAGEHNILPASTNP 302
QY 302 TIPTVNTVAEHLDMLTCHHLDKRREDLOFSQRIIPGSIATAEDVLHDMGVIAWTSDD 361
Db 303 TIPTVNTVAEHLDMLTCHHLDKRSIKEDVQFADSRIRPQTIAAEDTLHDMGIFITSDD 362
QY 362 SOAMGRAGEV1PRTWQTADKNKKEFGKLPEDGKONDNPRIKRYISKYKTINPALTHGVSEY 421
Db 363 SOAMGRVGEV1TRTWQTADKNKKEFGRLPEEKGDNDNFR1KRYISKYKTINPAITHGISEY 422
QY 422 IGSVEEGKIADLVVNPAPFGVKPKIV1KGGMVVFSMGDSNASVPTPOPVYVYREMFHGH 481

Db 423 VGSVEGKVADLVVNSPAFFGKPNMI1KGGFIALSQMGDANAS1PTPOPVYVYREMFHGH 482
QY 482 GKAKFDTSITFVSKVAYENGVEKLGRLERQVLVPVKNCRNITKQDFKNDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAYDKG1KEELGLERQVLVPVKNCRNITKQDMQFNDDTAHIEVNPET 542
QY 542 FEVVDGKLCSTKPTSQVPLAQRVTF 569
Db 543 YHVFVDGKEVTSKPNKVSQAQLFSIF 569
RESULT 3
US-08-467-822-26
; Sequence 26, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiherge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-822-26
Query Match 76.5%; Score 2294; DB 1; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.6e-40;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKKQEVNTYPTKGDVRLGDTDLMAVEVDHYTTYGEELKFGAGKTIREGMSQNSPD 61
Db 3 KISRKEVSMYGPPTGDKVRLGDTDLIAVEVDHYTYGEELKFGGKTIREGMSQNSPS 62

Db 3 KISRKEYVMVGPPTGDKVRLGDTDLIAEVEHDYTYIGBELKFGGKTLREGMSQSNPS 62
QY 62 ENTDLVITNAMIIDYTYKADIGIKNGKIHGKAGKMDQGVSPHVMVGVGTAL 121
Db 63 KEELDLIITNALIVDTYGIYKADIGIKNGKIHGKAGKMDQGVKNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHHFLSPQFPPTALANGVTMTFGGGTGPVDTGNATTITPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIFISPOQIPAPASGVMTMIGGGTGPADGTNATTITPGRRNLKWM 182
QY 182 LRAAEYSNMVGLFKNSSKKQLVEQVAGAIKFKLHEDWGTTPSAIDHCLSVAD 241
Db 183 LRAAEYSNMVGLFKNSSKKQLVEQVAGAIKFKLHEDWGTTPSAIDHCLSVAD 242
QY 242 VQVCHTDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHSPDVITMAGELNIP 301
Db 243 VQVAIHTDLNEAGCVEDTMAAIAAGTMTFTHTGAGGCHAPDIIKVAGEHNILP 302
QY 302 TIPTVINTVAEHLDMLTCHLDKRIREDLPQSRSIRPGSIAAEDVLHDMGVIA 361
Db 303 TIPTVTNEAEHMDMLVCHLDSIKEDVQPADSRIRPQTIAAEDTLHDMGIFS 362
QY 362 SQMGRAGEVIRPTWQADKNKEFGKLPEDGKNDNFRIKRYISKYTNIPALTH 421
Db 363 SQMGRVGEVITRTWQADKNKEFGKLPEDGKNDNFRIKRYISKYTNIPALTH 422
QY 422 IGSVEEGKIADLVNMPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPOP 481
Db 423 VGSVEGVKQVADLVNMPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPOP 482
QY 482 GKAKFDTSTITFVSKVAYENGVEKELGLERQVLPVKNCRNITKKDFKNDK 541
Db 483 GKAKYDRNITFVSOAAYDKGKEELGLERQVLPVKNCRNITKKDMQFNDT 542
QY 542 FEVFDGKLTCKTSKPTSQVPLAQRTPF 568
Db 543 YHVFVDGKVTSPANKVSLAQLFSIF 569

RESULT 4

US-08-432-697-26
; Sequence 26, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146

; REFERENCE/DOCKET NUMBER: 03495.0137-00000

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-432-697-26

Query Match 76.5%; Score 2294; DB 2; Length 569;

Best Local Similarity 73.4%; Pred. No. 1.6e-40;

Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMKKORYVNTYPTGDKVRLGDTDLIAEVEHDYTYIGBELKFGGKTLREGMSQSNPS 61
Db 3 KISRKEYVMVGPPTGDKVRLGDTDLIAEVEHDYTYIGBELKFGGKTLREGMSQSNPS 62
QY 62 ENTDLVITNAMIIDYTYKADIGIKNGKIHGKAGKMDQGVSPHVMVGVGTAL 121
Db 63 KEELDLIITNALIVDTYGIYKADIGIKNGKIHGKAGKMDQGVKNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHHFLSPQFPPTALANGVTMTFGGGTGPVDTGNATTITPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIFISPOQIPAPASGVMTMIGGGTGPADGTNATTITPGRRNLKWM 182
QY 182 LRAAEYSNMVGLFKNSSKKQLVEQVAGAIKFKLHEDWGTTPSAIDHCLSVAD 241
Db 183 LRAAEYSNMVGLFKNSSKKQLVEQVAGAIKFKLHEDWGTTPSAIDHCLSVAD 242
QY 242 VQVCHTDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHSPDVITMAGELNIP 301
Db 243 VQVAIHTDLNEAGCVEDTMAAIAAGTMTFTHTGAGGCHAPDIIKVAGEHNILP 302
QY 302 TIPTVINTVAEHLDMLTCHLDKRIREDLPQSRSIRPGSIAAEDVLHDMGVIA 361
Db 303 TIPTVTNEAEHMDMLVCHLDSIKEDVQPADSRIRPQTIAAEDTLHDMGIFS 362
QY 362 SQMGRAGEVIRPTWQADKNKEFGKLPEDGKNDNFRIKRYISKYTNIPALTH 421
Db 363 SQMGRVGEVITRTWQADKNKEFGKLPEDGKNDNFRIKRYISKYTNIPALTH 422
QY 422 IGSVEEGKIADLVNMPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPOP 481
Db 423 VGSVEGVKQVADLVNMPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPOP 482
QY 482 GKAKFDTSTITFVSKVAYENGVEKELGLERQVLPVKNCRNITKKDFKNDK 541
Db 483 GKAKYDRNITFVSOAAYDKGKEELGLERQVLPVKNCRNITKKDMQFNDT 542
QY 542 FEVFDGKLTCKTSKPTSQVPLAQRTPF 568
Db 543 YHVFVDGKVTSPANKVSLAQLFSIF 569

RESULT 5

US-08-466-248-26
; Sequence 26, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:


```

QY 242 VOVCHTDTVNEAGYVDDTLNMGRAIHAYHIEGAGGHSDDVITWAGELNLPSTTP 301
Db 243 VQVAIHTDTLNEAGCVEDTMAAIAGRTHWTFTEGAGGSHAPDIIKVAGEHNLPASTNP 302
QY 302 TTPVTINTVAEHLDMTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWSSD 361
Db 303 TTPFTVTEAEHMDMLVCHLDKSIKEDVQPADSRIRPOTIAAEDTLHDMGIFISITSSD 362
QY 362 SQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSSY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLKEKGGNDNFRIKRYISKYITINPAIAHGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPIVIGKGMVVFSEMGDSNASVPTPQPVYREMFHGH 481
Db 423 VGSVEVGKVDLVLSWPAFFGKPNMIKGGFIALSQMGDANASITPTPQPVYREMFPAHH 482
QY 482 GKAKFDTSTITFVSKVAYENGVEKGLERQVLVKNCRNITKCKPFENDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAAYDKGIEBELGLERQVLVKNCRNITKCKOMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKCLTSKPTSOVPLAQRY 565
Db 543 YHVFVDGKEVTSKANKVSLAQLP 566

```

```

RESULT 7
PCT-US96-05800-3
; Sequence 3, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: Oravax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-05800-3

```

Query Match 76.4%; Score 2290; DB 5; Length 566;

```

Best Local Similarity 73.6%; Pred. No. 2e-40;
Matches 415; Conservative 75; Mismatches 74; Indels 0; Gaps 0;
QY 2 KMKQEVNTYGTGDKVRLGDTDLWAVEHDYTYTYGELKFGAGKTIREGMQSQNSPD 61
Db 3 KISRKEYVSMYGTTGDKVRLGDTDLIAVEHDYTYTYGELKFGCGKTLREGMSQSNPS 62
QY 62 ENTLDLIVTNAMLIIDYTYGIKADIGIKNGKIHGKAGNKMDDGYSPHVVGVGTAL 121
Db 63 KEELDLIITNALIVDYGKADIGIKOGKIAGIGKGNKMDQGVKNLNSVGPATEALA 122
QY 122 GEGMIITAGIDSHFTLSPQOPPTALANGVTTMFGGTPGVDTNATTTTPGKNLHRM 181
Db 123 GGLIVTAGIIDIHIFISPOQIPTAFASGVITMGGTGPADGTNATTTTPGRNLUKM 182
QY 182 LRAAEYSNNVGLKGNSSSKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADEYD 241
Db 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIGFKIHEDWGTTPSAINHLDVADKYD 242
QY 242 VOVCHTDTVNEAGYVDDTLNMGRAIHAYHIEGAGGHSDDVITWAGELNLPSTTP 301
Db 243 VQVAIHTDTLNEAGCVEDTMAAIAGRTHWTFTEGAGGSHAPDIIKVAGEHNLPASTNP 302
QY 302 TTPVTINTVAEHLDMTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWSSD 361
Db 303 TTPFTVTEAEHMDMLVCHLDKSIKEDVQPADSRIRPOTIAAEDTLHDMGIFISITSSD 362
QY 362 SQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLKEKGGNDNFRIKRYISKYITINPAIAHGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPIVIGKGMVVFSEMGDSNASVPTPQPVYREMFHGH 481
Db 423 VGSVEVGKVDLVLSWPAFFGKPNMIKGGFIALSQMGDANASITPTPQPVYREMFPAHH 482
QY 482 GKAKFDTSTITFVSKVAYENGVEKGLERQVLVKNCRNITKCKPFENDKTAKITVDPKT 541
Db 483 GKAKYDANITFVSOAAYDKGIEBELGLERQVLVKNCRNITKCKOMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKCLTSKPTSOVPLAQRY 565
Db 543 YHVFVDGKEVTSKANKVSLAQLP 566

```

```

RESULT 8
US-09-431-705-5
; Sequence 5, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-431-705-5

```

```

Query Match 76.1%; Score 2282; DB 2; Length 570;
Best Local Similarity 73.0%; Pred. No. 2.9e-40;
Matches 416; Conservative 75; Mismatches 76; Indels 3; Gaps 1;
QY 2 KMKK---QEVNTYGTGDKVRLGDTDLWAVEHDYTYTYGELKFGAGKTIREGMQSQNS 58
Db 1 EMKKISRKEYVSMYGTTGDKVRLGDTDLIAVEHDYTYTYGELKFGCGKTLREGMSQSN 60

```

Qy		59	S P E N T L D L V I T W A M I I D Y G I K A D I G I K N K K I H G I K A G N Q K O M Q D G S P H M V G V G T E	118
Db		61	N P S K E E L D L I T N A I V I D Y T I G I K A D I G I K O G K I A G I K G K N K O M Q D G V X N L S V G P A T E	120
Qy		119	A L A G E M I I T A G G I D S R T H F L S P O O F P T A L A N G V T T M F G C G T P V D G T N A T T I P G K N L	178
Db		121	A L A G E G L I V T A G G I D T H I H F I S P O O I P T A F A S G V T T M I G G G T G P A D G T N A T T I T P G R N L	180
Qy		179	H R M L R A A B E Y S M N V G F L K G N S S K Q I V E Z A G A I G F K L H E D W G T T P S A I D H C L S V A D	238
Db		181	K W M L R A A B E Y S M N L G F L A K N A S N D A S L A D J E A G A I G F A H E D W G T T P S A I N H A L D V A D	240
Qy		239	E Y D V Q V C I H T D V N E A G Y V D T L N A M G R A I H A Y H I E G A G G H S P D V T I M A G E L N I L P S S	298
Db		241	K Y D V Q V A I A T D T L N E A G C V E D T M A A I A G R T W H T F H T E G A G G C H A P D I I K V A G E H N I L P A S	300
Qy		299	T T T I P T I N T V A E H L D M L M T C H H L D K R I R E D L O F S Q S R I P P G S I A A E D V L H D M G V I A M T	358
Db		301	T N T I P T V N T A E H M D M L A V C H H L D K S I K E D V Q F A D S R I R P Q T I A A E D T L H D M G I F S I T	360
Qy		359	S S D S Q A M G R A G E V I P R T W Q T A D N K N K E F G K L P E D G K O N D N F R I K R Y I S K Y T I N P A L T H G V	418
Db		361	S S D S Q A M G R V G E V I T R T W Q T A D N K N K S F G R U K E E K G O N D N F R I K R Y L S K Y T I N P A I A H G I	420
Qy		419	S E Y I G S V E E G K I A D L V V N P A F F G V G K R I V I K G M V F S E M G D S N A S V P T P Q P V Y R E M F	478
Db		421	S E Y V G S V E V G K V A D L V L W S P A F F G V K P N M I I K G G F I A L S Q M G D A N A S I P T P Q P V Y R E M F	480
Qy		479	G H G K A K F D T S I T P V S K V A Y E N G V K E K L G L E R Q V L P V K N C R N I T K Q D K F N D K T A K I T V D	538
Db		481	A H G K A K Y D A N I T P V S Q A A Y D K I K E E L G L E R Q V L P V K N C R N I T K Q D M Q F N D T T A H I E V N	540
Qy		539	P K T F E V F V D K L C T S K P T S Q V P L A Q R Y T F F	568
Db		541	P E T Y H V F V D G K E V T S K P A N K Y S L A Q L S I F	570

RESULT 9

```

US-08-467-822-21
; Sequence 21, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBLASTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-May-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..569
; OTHER INFORMATION: /note= "URE B - FIGURE 3."
US-08-432-697-21

Query Match          74.8%; Score 2244; DB 2; Length 569;
Best Local Similarity 72.5%; Pred. No. 1.8e-39;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;

QY      2 KMKCKEYVNTYGPTKGDVKRLGDTDLWAEVEHDVTYTGEEELKFAGAKTIRAGMGQSNSPD 61
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY     62 ENTLDLVITNAMIIDYGTGIYKADIGIKNGKIHGICKAGNKDMQDGVSPHVMVGTEALA 121
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY     63 SYVELDLVLTNALIVDYGTGIYKADIGIKDGKIAGICKAGNKDMQDGVNNLCVGPATEALA 122
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY    122 GGEMITAGGIDSHTHFSLSPQQFPALANGVTTFWGCGTGVDGTNATTTTPGKNLHRM 181
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY    123 ABEGLIVTAGGIDTHLFHSPPQIIPAFASGVTVTMGGGTGPAADGTNATTTTPGRANLKSM 182
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY    182 LRAAEYSNMVGFCLKGNSSKKQLAVEOVEAGAIFKLHEDDWGTPPSAIDHCLSVADEYD 241
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY    183 LRAAEYANNLUGFLKGNVSPEYPSLRDQIEAGAI GFKLHEDWGSTPRAIHHLNVDAEYD 242
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY    242 VQCVIHTDTVNEAGVVDDTLNAMNGRATHAYHI8GAGGGSHSPDVITWMAGELNILPSSSTP 301
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY    243 VQVAIHTDILNEAGCVETDLFAIAGRTIHTFHTEGACGGHAPDVIKWAGEFNILLPASTNP 302
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY    302 TPTYINTVAEHLMMLMTCHLDKRREDLOFSQRSRI RPGSIAAEDVLHMDGVAMTSSD 361
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY    303 TPFTFKNTAEAHMDMLMYCHLLDKSIKEDVOGFADFISRPQTIAAEOQLHWDGIFSITTSD 362
Db       | : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

QY	362	SOAMGRAGEVITRTWTQTADKNKKEFGKLPEBGGKNDNFRFKRYISKYTINPALTHGVSEY	421
D8	363	SOAMGRVEGVIITRTWTQTADKNKKEFGRLKEEKGDNNDNFRFKRYISKYTINPGIAHGISDY	422
QY	422	IGSVEEGKIADLVVNPAFPGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYYREMFGHH	481
D8	423	VGSVEBGKYADVLVWSPAFGLIPNMIIKGGFTALSGMDANASIPTPPQVVYYREMFGHH	482
QY	482	GKAKEDTTSITFYSKVAYENGVRKEKLGRLERQVLFPVKNCRNITKKDKFKENDKTAKITVDPKT	541
D8	483	GKNKFDTNITPVSQAAAYKAGIKEELGLDRAAPPVKNCRNITKKDLKENDVTAHIDVNPET	542
QY	542	FEVFDVGKLTCTSKPTSQVPLAQRYTFP	568
D8	543	YKRVYDGREVTSKADELSLAQLYNLF	569
RESULT 11			
US-08-466-248-21			
; Sequence 21, Application US/08466248			
; Patent No. 6258359			
; GENERAL INFORMATION:			
; APPLICANT: Labigne, Agnes			
; APPLICANT: Sauerbaum, Sebastien			
; APPLICANT: Ferrero, Richard L.			
; APPLICANT: Thiherge, Jean-Michel			
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST			
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE			
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID			
; TITLE OF INVENTION: POLYPEPTIDES			
; NUMBER OF SEQUENCES: 44			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &			
; ADDRESSEE: Dunner			
; STREET: 1300 I Street, N.W.			
; CITY: Washington			
; STATE: D.C.			
; COUNTRY: USA			
; ZIP: 20005-3315			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/466,248			
; FILING DATE: 06-JUN-1995			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/447,177			
; FILING DATE: 19-MAY-1995			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 08/432,697			
; FILING DATE: 02-MAY-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Meyers, Kenneth J.			
; REGISTRATION NUMBER: 25,146			
; REFERENCE/DOCKET NUMBER: 03495.0137-02000			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (202) 408-4000			
; TELEFAX: (202) 408-4400			
; INFORMATION FOR SEQ ID NO: 21:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 569 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: peptide			
; FEATURE:			
; NAME/KEY: Protein			

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26887

```

Query Match      61.2%; Score 1834; DB 2; Length 566;
Best Local Similarity 59.1%; Pred. No. 6.6e-31;
Matches 337; Conservative 87; Mismatches 140; Indels 6; Gaps 3;

Qy 1 MMKKQEVNTYGTGKDKVRLGDTLWAEVEHDYTTYGEELKFGAGKTIREGMGQSNSP 60
Db 1 MKISQAYADMFGPTVGRVRLADTLWIEVERDFTVYGEVVKFGGKVI RDGMGQSOLG 60

Qy 61 DENTLDLIVITNAMIIDYIYKADIGIKNGKIHGKAGNKMOMQDCVSPHVVGVGTEAL 120
Db 61 AAQVVDVTITNALILDHGWVVKADVGLKDGRIQAIKAGNPDIOFGV--NTAIGAGTEVI 118

Qy 121 AGEQMITAGGIDSHTHFLSPQPPALANGVYTMFGGGTGPVDTNATTITPGKWNLHR 180
Db 119 AGEQMITAGGIDSHTHFLSPQPPALANGVYTMFGGGTGPVDTNATTITPGKWNLHR 178

Qy 181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEBAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Db 179 MLQAADAPPNNIGFTGKGNASLPLPLEBQVLAGAIGLKHEDWGSTPAADNCLLEVAERH 238

Qy 241 DVQVCIHTDVTNEAGYVDDTLNANNGRAIHAYHIEGAGGHSPPVITWAGEINILPSSTT 300
Db 239 DIQVAIHTDTLNEGFEVETTLGAFKGRITHTYHTGAGGHPADIIKACGFANVLPSSTN 298

Qy 301 PTIPVTINTVAEHLMDLMTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSS 360
Db 299 PTRPPTRTNIDHLDMLVCHHLDPAIEDVAFAPESRIRRETIAAEDILHDLGAFSMISS 358

Qy 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPBGDKNDNFRIKRYISKYTIINPALTHGVSE 420
Db 359 DSQAMGRAGEVITRTWQTADKNKQGRGLDGDGAENDNFRARRYIAKYTIINPAITHGISH 418

Qy 421 YIGSVBEGKIADLVVWNPAPFGVPKIVIKGGMVVFSEMGDSNASVPTPPQVYVREMFGH 480
Db 419 EVGVSVEAGKWADLVLRPAFFGVKPSLILKGAIAAASLMGDIINGSIPTPPQVHYRPMFAS 478

Qy 481 HGKAKFDTISITFVSKVAYENGKVKGLERQVLVYKNCNITKKDFKENDKTAKITVDPK 540
Db 479 YAGSRHATSLTFVSOAFAAGVPOQLGRKAIGVVSVCGRGVQKTDLIHNGYLPTEIYDAQ 538

Qy 541 TFEVFDGKL--CTSKPTSQVPLAORYTFF 568
Db 539 NYQVRADGQLLWC--EPADVLPMAORYFLP 566

```

Search completed: August 9, 2006, 21:44:42
Job time : 53 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:44:02 ; Search time 185 Seconds

(without alignments)
1422.195 Million cell updates/sec

Title: US-09-904-994B-3

Perfect score: 2999

Sequence: 1 MMKKQEVNTYGTGDKV.....KLCTSKPTSQVPLQRYTFF 568

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main.*

- 1: /EMC_Celerra_SID3S/prodata/2/pubpaa/US07_PUBCOMB.psp.*
- 2: /EMC_Celerra_SID3S/prodata/2/pubpaa/US08_PUBCOMB.psp.*
- 3: /EMC_Celerra_SID3S/prodata/2/pubpaa/US09_PUBCOMB.psp.*
- 4: /EMC_Celerra_SID3S/prodata/2/pubpaa/US10A_PUBCOMB.psp.*
- 5: /EMC_Celerra_SID3S/prodata/2/pubpaa/US10B_PUBCOMB.psp.*
- 6: /EMC_Celerra_SID3S/prodata/2/pubpaa/US11_PUBCOMB.psp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2999	100.0	568	3	US-09-904-994B-3
2	2973	99.1	568	3	US-09-904-994B-15
3	2960	98.7	568	3	US-09-904-994B-6
4	2953	98.5	568	3	US-09-904-994B-12
5	2591	86.4	496	3	US-09-904-994B-9
6	2326	77.6	569	4	US-10-639-273-5
7	2304	76.8	568	4	US-10-639-273-38
8	2297	76.6	569	3	US-09-895-913A-252
9	2297	76.6	569	4	US-10-282-122A-58611
10	2297	76.6	569	4	US-10-335-977-8611
11	2297	76.6	569	4	US-10-639-273-40
12	2244	74.8	569	4	US-10-639-273-39
13	2239	74.7	750	3	US-09-402-100-2
14	2109	70.3	569	4	US-10-639-273-41
15	1927	64.3	568	4	US-10-282-122A-47638
16	1921	64.1	566	4	US-10-282-122A-69286
17	1913	63.8	568	4	US-10-282-122A-49766
18	1902	63.4	568	4	US-10-282-122A-50238
19	1843	61.5	572	4	US-10-282-122A-58210
20	1838	61.3	566	4	US-10-282-122A-66743
21	1834	61.2	567	4	US-10-282-122A-60007
22	1831	61.1	779	4	US-10-282-122A-69134
23	1826	60.9	566	4	US-10-282-122A-44964
24	1824	60.8	837	4	US-10-424-599-254635
25	1818	60.6	568	4	US-10-282-122A-56823
26	1818	60.6	568	4	US-10-282-122A-56844
27	1810	60.4	567	4	US-10-282-122A-68252

28	1804	60.2	571	5	US-10-793-626-118	Sequence 118, App
29	1804	60.2	811	4	US-10-282-122A-70574	Sequence 70574, A
30	1803	60.1	573	4	US-10-724-972A-5526	Sequence 5526, Ap
31	1797	59.9	571	4	US-10-282-122A-44443	Sequence 44443, A
32	1795	59.9	572	4	US-10-282-122A-78027	Sequence 78027, A
33	1770	59.0	840	4	US-10-621-833-7	Sequence 7, Appli
34	1770	59.0	840	5	US-10-731-877-1	Sequence 1, Appli
35	1770	59.0	840	6	US-11-046-271-1	Sequence 1, Appli
36	1769	59.0	843	6	US-10-437-963-124057	Sequence 124057, A
37	1728	57.6	870	3	US-09-738-626-3599	Sequence 3599, Ap
38	1726	57.6	872	4	US-10-418-962-2	Sequence 2, Appli
39	1721	57.4	598	4	US-10-282-122A-76849	Sequence 76849, A
40	1720	57.4	570	5	US-10-454-437-14	Sequence 14, Appl
41	1713	57.1	571	4	US-10-282-122A-51072	Sequence 51072, A
42	1706	56.9	573	4	US-10-156-761-14635	Sequence 14635, A
43	1631	54.4	534	5	US-10-454-437-16	Sequence 16, Appl
44	1624	54.2	577	4	US-10-282-122A-62532	Sequence 62532, A
45	1624	54.2	577	4	US-10-282-122A-64617	Sequence 64617, A

ALIGNMENTS

RESULT 1
US-09-904-994B-3
; Sequence 3, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1e1 N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Depoedesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-3

Query Match 100.0%; Score 2999; DB 3; Length 568;
Best Local Similarity 100.0%; Pred. No. 2.3e-53;
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTTTGEELKFGAGKTIREGMGQSNP	60
Db	1	MMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTTTGEELKFGAGKTIREGMGQSNP	60
Qy	61	DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSPhmVVGVGTEAL	120
Db	61	DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSPhmVVGVGTEAL	120
Qy	121	AGEGMIITAGGIDSHTHPLSQOFPPTALANGVTWTFGGTGVDTGNATTITPGKNLHR	180
Db	121	AGEGMIITAGGIDSHTHPLSQOFPPTALANGVTWTFGGTGVDTGNATTITPGKNLHR	180
Qy	181	MLRAEYSNMVFLGKGNSSKKQLVQVEAGAGIKLHEDMGTTSPSIDHCLSVADY	240
Db	181	MLRAEYSNMVFLGKGNSSKKQLVQVEAGAGIKLHEDMGTTSPSIDHCLSVADY	240
Qy	241	DVQVCIHDTVNEAGYVDDTLNMGRAIHAYHIEGAGGGHSPDVTWAGELNLPSSIT	300
Db	241	DVQVCIHDTVNEAGYVDDTLNMGRAIHAYHIEGAGGGHSPDVTWAGELNLPSSIT	300
Qy	301	PTIPTYINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEVDLHDMGVIAVTSS	360
Db	301	PTIPTYINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEVDLHDMGVIAVTSS	360
Qy	361	DSQAMGAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTGHVSE	420
Db	361	DSQAMGAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTGHVSE	420

Qy 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Db 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Qy 481 HGKAKFDTSTFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKENDTKAKITVDPK 540
Db 481 HGKAKFDTSTFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKENDTKAKITVDPK 540
Qy 541 TFEVFDGKCTSKPTSQVPLAQRVTF 568
Db 541 TFEVFDGKCTSKPTSQVPLAQRVTF 568

RESULT 2
US-09-904-994B-15
; Sequence 15, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-15

Query Match 99.1%; Score 2973; DB 3; Length 568;
Best Local Similarity 98.9%; Pred. No. 7.9e-53;
Matches 562; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60
Db 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60
Qy 61 DENTLDLVITNAMIIDYTGIIKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Db 61 DENTLDLVITNAMIIDYTGIIKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Qy 121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLHR 180
Db 121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLHR 180
Qy 181 MLRAAEYSMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Db 181 MLRAAEYSMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Qy 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNILPSSTT 300
Db 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNILPSSTT 300
Qy 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Db 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Qy 361 DSQAMGRAGEVTPRTWQTADKNKKEFGKLPEGDGKNDNFRIKRYISKYITINPALTHGVSE 420
Db 361 DSQAMGRAGEVTPRTWQTADKNKKEFGKLPEGDGKNDNFRIKRYISKYITINPALTHGVSE 420
Qy 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Db 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Qy 481 HGKAKFDTSTFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKENDTKAKITVDPK 540
Db 481 HGKAKFDTSTFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKENDTKAKITVDPK 540
Qy 541 TFEVFDGKCTSKPTSQVPLAQRVTF 568
Db 541 TFEVFDGKCTSKPTSQVPLAQRVTF 568

RESULT 4
US-09-904-994B-12
; Sequence 12, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine

Db 541 TFEVFDGKCTSKPTSEVPLAQRVTF 568

RESULT 3
US-09-904-994B-6
; Sequence 6, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-6

Query Match 98.7%; Score 2960; DB 3; Length 568;
Best Local Similarity 98.6%; Pred. No. 1.5e-52;
Matches 560; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60
Db 1 MMKKQEVNTYGTGDKVRLGDTDLMAEVEHDYTYGEEELKFGAGKTIREGMQSNP 60
Qy 61 DENTLDLVITNAMIIDYTGIIKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Db 61 DENTLDLVITNAMIIDYTGIIKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTEAL 120
Qy 121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLHR 180
Db 121 AGEGLIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLHR 180
Qy 181 MLRAAEYSMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Db 181 MLRAAEYSMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
Qy 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNILPSSTT 300
Db 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGELNILPSSTT 300
Qy 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Db 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
Qy 361 DSQAMGRAGEVTPRTWQTADKNKKEFGKLPEGDGKNDNFRIKRYISKYITINPALTHGVSE 420
Db 361 DSQAMGRAGEVTPRTWQTADKNKKEFGKLPEGDGKNDNFRIKRYISKYITINPALTHGVSE 420
Qy 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Db 421 YIGSVEGKIADLVVWNPAPFGVGPVKIVIKGMVVFSEMGDSNASVPTPQPVYIREMFGH 480
Qy 481 HGKAKFDTSTFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKENDTKAKITVDPK 540
Db 481 HGKAKFDTSTFVSKVAVENGKGLGLERQVLPVKNCRNITKKDFKENDTKAKITVDPK 540
Qy 541 TFEVFDGKCTSKPTSQVPLAQRVTF 568
Db 541 TFEVFDGKCTSKPASEVPLAQRVTF 568

RESULT 4
US-09-904-994B-12
; Sequence 12, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine

FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-12

Query Match 98.5%; Score 2953; DB 3; Length 568;
Best Local Similarity 98.2%; Pred. No. 2e-52; Mismatches 5; Indels 0; Gaps 0;
Matches 558; Conservative 5;
QY 1 MKMKQEVYNTYPTGDKVRLGDTDLWAEVEHDYTYTYGELKFGAGKTIREGNGQSNP 60
DB 1 MKMKQEVYNTYPTGDKVRLGDTDLWAEVEHDYTYTYGELKFGAGKTIREGNGQSNP 60
QY 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSHPMVVGVGTEAL 120
DB 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSHPMVVGVGTEAL 120
QY 121 AGEEMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDTGNATTITPGKNLHR 180
DB 121 AGEEMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDTGNATTITPGKNLHR 180
QY 181 MLRAAEYSNMVGFGLKGNSSKKQLVEQVEAGAIGFKLHEDMGTTTSAIDHCLSLVADEY 240
DB 181 MLRAAEYSNMVGFGLKGNSSKKQLVEQVEAGAIGFKLHEDMGTTTSAIDHCLSLVADEY 240
QY 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTT 300
DB 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTT 300
QY 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
DB 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
QY 361 DSQAMGRAGEVLPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
DB 361 DSQAMGRAGEVLPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
QY 421 YIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYYREMFGEH 480
DB 421 YIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYYREMFGEH 480
QY 481 HGKAKFDTSTITFVSQVAYENGVEKGLERQVLPVKNCRNITKKDPFNKTAKITVDPK 540
DB 481 HGKAKFDTSTITFVSQVAYENGVEKGLERQVLPVKNCRNITKKDPFNKTAKITVDPK 540
QY 541 TFEVFDGKLTCKTSPKASEVPLAQRYTFF 568
DB 541 TFEVFDGKLTCKTSPKASEVPLAQRYTFF 568

RESULT 5
US-09-904-994B-9
; Sequence 9, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-9

Query Match 86.4%; Score 2591; DB 3; Length 496;
Best Local Similarity 99.2%; Pred. No. 4.2e-45; Mismatches 2; Indels 1; Gaps 1;
Matches 491; Conservative 1;
QY 1 MKMKQEVYNTYPTGDKVRLGDTDLWAEVEHDYTYTYGELKFGAGKTIREGNGQSNP 60
DB 1 MKMKQEVYNTYPTGDKVRLGDTDLWAEVEHDYTYTYGELKFGAGKTIREGNGQSNP 60
QY 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSHPMVVGVGTEAL 120
DB 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSHPMVVGVGTEAL 120
QY 121 AGEEMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDTGNATTITPGKNLHR 180
DB 121 AGEEMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGPVDTGNATTITPGKNLHR 180
QY 181 MLRAAEYSNMVGFGLKGNSSKKQLVEQVEAGAIGFKLHEDMGTTTSAIDHCLSLVADEY 240
DB 181 MLRAAEYSNMVGFGLKGNSSKKQLVEQVEAGAIGFKLHEDMGTTTSAIDHCLSLVADEY 240
QY 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTT 300
DB 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTMAGELNLPSTT 300
QY 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
DB 301 PTIPYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVTSS 360
QY 361 DSQAMGRAGEVLPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
DB 361 DSQAMGRAGEVLPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
QY 421 YIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYYREMFGEH 480
DB 421 YIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYYREMFGEH 480
QY 481 HGKAKFDTSTITFVS 494
DB 481 HGKAKFDTSTITFRVS 495

RESULT 6
US-10-639-273-5
; Sequence 5, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 5
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-5

Query Match 77.6%; Score 2326; DB 4; Length 569;
Best Local Similarity 74.1%; Pred. No. 1.5e-39;
Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;
QY 2 KMKKQEVYNTYPTGDKVRLGDTDLWAEVEHDYTYTYGELKFGAGKTIREGNGQSNP 61
DB 3 KISRKEYVSMYPTGDKVRLGDTDLILEVEHDCYTYGEEIKFPGGKTIIDGMAQTNSPS 62

Qy	62	ENTLDLVITNAMIIDYTGIIYKADIGIKNGKIHIGIKAGNKMQDQGVSHMVVGVGTEALA	121
Db	63	SHELDLVITNALIYDYGIIYKADIGIKNGKIHIGIKAGNKMQDQGVNCLNCGVPATEALA	122
Qy	122	GEGMITAGGIDSHTFLSPQOFPTALANGVTWTFMGCGTGPVDGNTATITPGKNLHRM	181
Db	123	ABGLIVTAGGIDTHIFISPOOIPTAFASGITWIMGGGTGPADGNTATITPGRWNLKTM	182
Qy	182	LRAABEYSNMVFLGKGNSSKKQIUVQVEAGAIQFKLHEDWGTTPSAIDHCLSVADEYD	241
Db	183	LRASEEYAMNLGYLGKGNVSEPSLVDQLEAGIQFKLHEDWGSTPAIYHCLNVADKYD	242
Qy	242	VQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSFVDVTMAGELNLPSTTP	301
Db	243	VQVAIHDTTNEAGCVEDTLQAIAGRTIHTPHTEGAGGCHAPDVIKMSGEFNILPASTNP	302
Qy	302	TIPYTIINTVAEHLDMLTCHHLDKRIREDLOFSQSRIRPGSIAABDVLDHMGVIAMTSSD	361
Db	303	TIPFTVNTAEHMDMLVCHHLDKNIKEDVOFADSRIRPQTIAABDKLHDMGIFITSDD	362
Qy	362	SQAMGRAGEVIPRTWOTADNKKRFGKLPEDGKONDNFRIKRYISKYITNPALTHGVSEY	421
Db	363	SQAMGRVGEVITRTWQADNKKRFGRLPEBKGDNDNFRIKRYISKYITNPAIAHIGISEY	422
Qy	422	IGSVEEGKIADLVVNPAFFGVKPKIVITKGVVVPSEMGDSNASVPTQPVYVRMFQGH	481
Db	423	VGSVEVGKFADLVLSPAFFGIKPMWIIKGGFIALSQMGDANASITPQPVYVRMFQGH	482
Qy	482	GKAKFDTSITFVSKVAYENGVEKLGLEBROVLVPVKNCRNITKKOPKFNDDTKAKITVDPKT	541
Db	483	GKAKFDNITFVSVQVADNGIKEELGLQVRVLVPVKNCRNITKKDLKFNDDVTAHIEVNPT	542
Qy	542	FEVVDGKLCTSKPQTSQVPLAQRYTFF	568
Db	543	YKVKVDGKEVTSKAADKISLAQLYNLF	569

```

RESULT 7
US-10-639-273-38
; Sequence 38, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 568
; TYPE: PR1
; ORGANISM: Helicobacter heilmannii
US-10-639-273-38

```

		Query Match	76.8%; Score 2304; DB 4; Length 568;
		Best Local Similarity	73.7%; Pred. No. 4.le-39;
		Matches	418; Conservative 74; Mismatches 74; Indels 1; Gaps 1;
Qy	2	KMKQEVVNTYGPTKDVKRLGDTDLWAEVEVDHYTTTGEELKFAGKTIIRSGMGGSNSPD	61
Dd	:	: : : : :	:
Dd	3	KISRKEYSVYGPTTGDKVKRLGDTDLLEVEHDCTTYGEEIKFGGGKTIRDGMGTNPS	62
Qy	62	ENTLDLVTNWMIIDYGIKYADIGIKNGKTHGICAKGNKDMDQGVSPHMYVGVTAL	121
Dd	:	: : : : :	:
Dd	63	SHELDLVITNALIIVDYTIKYADIGIKNGKTHGICAKGNKDLQDGVCNLCPGPAATEALA	122
Ov	122	GEGMIIITAGGISDHTFLSPFOFPITALANGVTWTMFGGGTGPVDGTNAITITIPGWKNLRHM	181

[illegible]

RESULT 8

```

US-09-895-913A-252
; Sequence 252, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Poly
; TITLE OF INVENTION: Encoding No. US200201
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-252

```

	Query Match	76.6%	Score 2297;	DB 3;	Length 569;
	Best Local Similarity	73.4%;	Pred. No. 5.7e-39;		
	Matches 416;	Conservative 76;	Mismatches 75;	Indels 0;	Gaps 0
Qy	2	KMKQEVYNTVYPTGDKVRLGDTDLAEVVEHDYTYGEEELKFCAGKTIIRGMSGQNSPD	61		
Db	3	KISRKEVSMYPTGDKVRLGDTDLAEVVEHDYTYGEEELKFCGKTLREGMSGQNSNPS	62		
Qy	62	ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKQMDQGVSPHMYVGVTEALA	121		
Db	63	KEELDILITNALIVDYTYGIYKADIGIKDGKIAGIKGKGNKMDQGVKKNLSVGPATEALA	122		
Qy	122	GEGMIITAGGIDSHTFLSPQFPFALANGVITNMFPGGTGPVDGNTNATITIPGKNLHRM	181		
Db	123	GEGLIVTAGGIDTHIFTSPOIETPASGVYTIMGGTGPDAGDTNATITIPGRNLKWM	182		


```
QY 182 LRAAEYSNMVGLKGNSSKKQLVQVEAGATGPKLHEDWGTTPSAIDHCLSVADKYD 241
D 183 LRAAEYSNMVGLKGNSSKKQLVQVEAGATGPKLHEDWGTTPSAIDHCLSVADKYD 242
QY 242 VQVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHGSPDVTWAGELNIPSTTP 301
D 243 VQVAIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHGSPDVTWAGELNIPSTTP 302
QY 302 TIPTVINTVAEHLMLTCHHLDKRIREDLQFSQSRIRPGSIAEDVLHDMGVAMTSSD 361
D 303 TIPTVNTAEAHMDMLVCHHLDKSIEDVQFADSRIRPQTIAAEDTLHDMGIFSISSD 362
QY 362 SOAMGRAGEVITPTWQADKNKKEFGKLPEDCKNDNFRIRKYLKYINPAIHGISEY 421
D 363 SOAMGRAGEVITPTWQADKNKKEFGKLPEDCKNDNFRIRKYLKYINPAIHGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQVYVREMFHGH 481
D 423 VGSVEGKADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQVYVREMFHGH 482
QY 482 GKAKFDTSIPTVSKVAYENGVEKELGLERQVLVKNKCRNITKKDFKFNOKTAKITVDPKT 541
D 483 GKAKYDANITFVSOAAYDKGKEELGLERQVLVKNKCRNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCSTKPTSQVPLAQRYTFF 568
D 543 YHVFVDGKEVTSKPANKVSLAQFSIF 569

RESULT 9
US-10-282-122A-58611
; Sequence 58611, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zysek, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58611
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-282-122A-58611
```

```
Query Match 76.6%; Score 2297; DB 4; Length 569;
Best Local Similarity 73.4%; Pred. No. 5.7e-39;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY 2 KMKQBYVNTYPTGDKVRLGDTDLWAEVEHDYTYTGELKEFGAGKTIREGMSQNSPD 61
D 3 KISRKEYVSMYPTGDKVRLGDTDLWAEVEHDYTYTGELKEFGAGKTIREGMSQNSPD 62
QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGNKMDDGVSPHVVGVGTBALA 121
D 63 KEELDLITNAILVDYTYGIYKADIGIKDKGIAGIKGKNKMDGVKNLNVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGCTGPDVGTNATTTPKWNLHRM 181
D 123 GEGLIIVTAGGIDTHIFISPOIPTAFASGVTTMIGGCTGPDVGTNATTTPGRRNLKWM 182
QY 182 LRAAEYSNMVGLKGNSSKKQLVQVEAGATGPKLHEDWGTTPSAIDHCLSVADKYD 241
D 183 LRAAEYSNMVGLKGNSSKKQLVQVEAGATGPKLHEDWGTTPSAIDHCLSVADKYD 242
QY 242 VQVCIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHGSPDVTWAGELNIPSTTP 301
D 243 VQVAIHDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGCHGSPDVTWAGELNIPSTTP 302
QY 302 TIPTVINTVAEHLMLTCHHLDKRIREDLQFSQSRIRPGSIAEDVLHDMGVAMTSSD 361
D 303 TIPTVNTAEAHMDMLVCHHLDKSIEDVQFADSRIRPQTIAAEDTLHDMGIFSISSD 362
QY 362 SOAMGRAGEVITPTWQADKNKKEFGKLPEDCKNDNFRIRKYLKYINPAIHGISEY 421
D 363 SOAMGRAGEVITPTWQADKNKKEFGKLPEDCKNDNFRIRKYLKYINPAIHGISEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQVYVREMFHGH 481
D 423 VGSVEGKADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQVYVREMFHGH 482
QY 482 GKAKFDTSIPTVSKVAYENGVEKELGLERQVLVKNKCRNITKKDFKFNOKTAKITVDPKT 541
D 483 GKAKYDANITFVSOAAYDKGKEELGLERQVLVKNKCRNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCSTKPTSQVPLAQRYTFF 568
D 543 YHVFVDGKEVTSKPANKVSLAQFSIF 569
```

```
RESULT 10
US-10-335-977-8611
; Sequence 8611, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
```

```
;; SOFTWARE: UNIX
;; CURRENT APPLICATION DATA:
;;   APPLICATION NUMBER: US/10/335,977
;;   FILING DATE: 30-Dec-2002
;; PRIOR APPLICATION DATA:
;;   APPLICATION NUMBER: 08/993,002
;;   FILING DATE: 17-DEC-1997
;; ATTORNEY/AGENT INFORMATION:
;;   NAME: Mandragouras, Amy E.
;;   REGISTRATION NUMBER: 36,207
;;   REFERENCE/DOCKET NUMBER: GTN-018
;; TELECOMMUNICATION INFORMATION:
;;   TELEPHONE: (617)227-7400
;;   TELEFAX: (617)742-4214
;; INFORMATION FOR SEQ ID NO: 8611:
;;   SEQUENCE CHARACTERISTICS:
;;     LENGTH: 569 amino acids
;;     TYPE: amino acid
;;     TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHEetical: YES
;; ORIGINAL SOURCE:
;;   ORGANISM: Helicobacter pylori
;; FEATURE:
;;   NAME/KEY: misc feature
;;   LOCATION: (B) LOCATION 1...569
;;   SEQUENCE DESCRIPTION: SEQ ID NO: 8611:
US-10-335-977-8611

Query Match      76.6%; Score 2297; DB 4; Length 569;
Best Local Similarity 73.4%; Pred. No. 5.7e-39;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY      2 KMKQKQEVNTYGTGKDVRLGDTDLWAWEVDHYTTYGBELKFGAGKTIREGMGQSNPSD 61
DB      3 KISRKEYVSMYGTPTGDKVRLGDTDLIAEVEHDYTYGBELKFGGKTLREGMQSNPS 62
QY      62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKQMDQGVSPHVVGVGTEALA 121
DB      63 KEELDLIITNALIVDYTYGIYKADIGIKGKIAGIKGKGNKQMDQGVKNLSVGPATEALA 122
QY      122 GEGMIITAGGIDSHYFLSPQPFPTALANGVVTMTFGGGTGPVDGTNATTITPGKNNLHRM 181
DB      123 GEGLVITAGGIDTHIFISPOQIPTAFASGVVTMTIGGGTGPADGTNATTITPGRNLLKM 182
QY      182 LRAAEYSNMVGLGKGNSSKKQIYVEQVGAIGFKLHEDMGTTPSAIDHCLSVADVD 241
DB      183 LRAAEYSNMVGLGKGNSSKKQIYVEQVGAIGFKLHEDMGTTPSAIDHCLSVADVD 242
QY      242 VQVCITHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNIPSTTP 301
DB      243 VOVAIHTDTLNEAGVEDTMAAIGRTWHTFHTGAGGGHAPDIIVKAGEHNILPASTNP 302
QY      302 TIPTVTINTVAEHLDMMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSD 361
DB      303 TIPTVTNTAEHMDMLVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSISSD 362
QY      362 SOAMGRAGEVITPRWTQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTGHVSEY 421
DB      363 SOAMGRVGEVITRTWTQADKNKKEFGRLKEEGKNDNFRIKRYISKYITINPAIAHGISEY 422
QY      422 IGSVEEGKIADLVVNPAPFFGVKPKIVIKGGWVVFSEMGDSNASVPTPOPVYVYREMPFGH 481
DB      423 VGSVEVGKADLVLSVPAFFGVKPNMIIKGGFIALSQMGDANASITPTPOPVYVYREMPFAH 482
QY      482 GKAKFDTSITFVSKVAYENGKVKGLERQVLVPRKCRNITKKDKFPNDKTKAKITVDPKT 541
DB      483 GKAKYDANITFVSAAYDKGIKEELGLERQVLVPRKCRNITKKDMQFNDTTAHIEVNPET 542
QY      542 FEVFDGKLCSTKPTSOVPLAQRYTFP 568
DB      543 YHVFVDGKEVTSKPANKVSLAQFLSFIF 569
```

RESULT 12

US-10-639-273-39

; Sequence 39, Application US/10639273

; Publication No. US20040142343A1

```
RESULT 11
US-10-639-273-40
; Sequence 40, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 569
; TYPE: PRP
; ORGANISM: Helicobacter pylori
US-10-639-273-40
```

```
Query Match      76.6%; Score 2297; DB 4; Length 569;
Best Local Similarity 73.4%; Pred. No. 5.7e-39;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

QY      2 KMKQKQEVNTYGTGKDVRLGDTDLWAWEVDHYTTYGBELKFGAGKTIREGMGQSNPSD 61
DB      3 KISRKEYVSMYGTPTGDKVRLGDTDLIAEVEHDYTYGBELKFGGKTLREGMQSNPS 62
QY      62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKQMDQGVSPHVVGVGTEALA 121
DB      63 KEELDLIITNALIVDYTYGIYKADIGIKGKIAGIKGKGNKQMDQGVKNLSVGPATEALA 122
QY      122 GEGMIITAGGIDSHYFLSPQPFPTALANGVVTMTFGGGTGPVDGTNATTITPGKNNLHRM 181
DB      123 GEGLVITAGGIDTHIFISPOQIPTAFASGVVTMTIGGGTGPADGTNATTITPGRNLLKM 182
QY      182 LRAAEYSNMVGLGKGNSSKKQIYVEQVGAIGFKLHEDMGTTPSAIDHCLSVADVD 241
DB      183 LRAAEYSNMVGLGKGNSSKKQIYVEQVGAIGFKLHEDMGTTPSAIDHCLSVADVD 242
QY      242 VQVCITHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNIPSTTP 301
DB      243 VOVAIHTDTLNEAGVEDTMAAIGRTWHTFHTGAGGGHAPDIIVKAGEHNILPASTNP 302
QY      302 TIPTVTINTVAEHLDMMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSD 361
DB      303 TIPTVTNTAEHMDMLVCHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSISSD 362
QY      362 SOAMGRAGEVITPRWTQADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTGHVSEY 421
DB      363 SOAMGRVGEVITRTWTQADKNKKEFGRLKEEGKNDNFRIKRYISKYITINPAIAHGISEY 422
QY      422 IGSVEEGKIADLVVNPAPFFGVKPKIVIKGGWVVFSEMGDSNASVPTPOPVYVYREMPFGH 481
DB      423 VGSVEVGKADLVLSVPAFFGVKPNMIIKGGFIALSQMGDANASITPTPOPVYVYREMPFAH 482
QY      482 GKAKFDTSITFVSKVAYENGKVKGLERQVLVPRKCRNITKKDKFPNDKTKAKITVDPKT 541
DB      483 GKAKYDANITFVSAAYDKGIKEELGLERQVLVPRKCRNITKKDMQFNDTTAHIEVNPET 542
QY      542 FEVFDGKLCSTKPTSOVPLAQRYTFP 568
DB      543 YHVFVDGKEVTSKPANKVSLAQFLSFIF 569
```


; APPLICANT: Zhu, Jiaqian
 ; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
 ; FILE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
 ; FILE REFERENCE: 19603/3881
 ; CURRENT APPLICATION NUMBER: US/10/639,273
 ; CURRENT FILING DATE: 2003-08-12
 ; PRIOR APPLICATION NUMBER: 60/404,337
 ; PRIOR FILING DATE: 2002-08-16
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 41
 ; LENGTH: 569
 ; TYPE: PRT
 ; ORGANISM: Helicobacter hepaticus
 ; US-10-639-273--41

Query Match 70.3% Score 2109; DB 4; Length 569;
 Best Local Similarity 68.1%; Pred No. 4.le-35;
 Matches 388; Conservative 83; Mismatches 95; Indels 4; Gaps 4;

QY	1	MMKKQOEYVNTYGTPTGDKVRLGDTDLWAEVEHYDYYTYGEBELFGAGKTIKREMGQSNP	60
DB	2	IKISRQYASMYGPTTGGDKVRLGDTNLFAELKDYDLYGEEIKFGGKTIKRDGMAQSAST	61
QY	61	DENTLDVITNAMIIDYTGTYKADIGIKNGKIRKIGKAGKMDQGVSPHMVGVGTEAL	120
DB	62	YTNELDAVITNAMIIDYTGTYKADIGIKGGKIKVIGKAGNPDTQDSVNEAMVVGAATEVI	121
QY	121	AGRGMIITAGGIDSHFTLSPPQPTTALANGVTTMFGSGTGPVDGTNATITTPCKWNLHR	180
DB	122	AGRGQIITAGGIDTHIHFIPTQIPTALYSGVVTMIGGGTGPAAGTNATCTPGKWNHQ	181
QY	181	MLRAAEYSNMNVGFLGKGNSSKKQLVEQVEACAIGFKLHEDWCTTTPSAIDHCLLSVADEY	240
DB	182	MLRAAESYANVLGFFGKGNSSNEEGLESEQIKAGALGLKVHEDWGSTPAAINHALNVAQKY	241
QY	241	DVQVCIHTDFTVNAGYVDDTLTNAMNGRAIHAYHI EGAGGGHSPDVI TWAGELNLPSSTT	300
DB	242	DVQVAIHTDFTLNAGCVEDTWKADIGRTIHTFTEGAGGGHAPDIIKAAAGEPNILPASTN	301
QY	301	PTTPYTINTVAEHLDMLTCHILDKRIFREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSS	360
DB	302	PTTPFTKNTADEHLDMLVMVCHLDKKIKEDVAFADSRIRPETIAAEDTLHDMGIFSITSS	361
QY	361	DSQAMGRAGREVIPTRTWOTADKNKKEFGKLPED - GKNDNDFRIKRYISKYITINPALTHGVS	419
DB	362	DSQAMGRGVEVITRTWOTADKNKKEFGALKEEG - ENDNDFRIKRYISKYITINPALHAGIS	420
QY	420	EYIGSVBEGKIADLVVNPAFAGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMF	479
DB	421	EYVGSVEVGKPADLVLMKPSMFGIKPEMILKNGMIVAAKIGDSNASIPTPEPVVYAPMFG	480
QY	480	HGKAFDTSITFVSKVAYENGVEKILGLERQVLVKNCRNITKKDKFNDKTKAITVDP	539
DB	481	SYGKAKYNCAITFVSKTAYDCHIKKEELGLERILLPVKNCRNITKKDKFNDVITPIEVNP	540
QY	540	KTFEVEFVDG - KLCTSKPTSOVPLAQRYTEF	568
DB	541	ETFEVRVNNTKI - TSKPEVKVSLGQLCLP	569

RESULT 15
 US-10-282-122A-47638
 ; Sequence 47638, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Hasebeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel

```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47638
; LENGTH: 568
; TYPE: PRF
; ORGANISM: Burkholderia cepacia
; US-10-282-122A-47638

Query Match          64.3%; Score 1927; DB 4; Length 568;
Best Local Similarity 61.3%; Pred. No. 2.2e-31;
Matches 348; Conservative 97; Mismatches 121; Indels 2; Gaps 1

Qy 1 MMKKQBYVNTYGTGDKVRLGDTDLWAEVHDYTYTYGEEKFGAGKTIREGMGQSNP 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LRLSRRAYAEFGPTGDRVRLADTELLIETIERDFTTYGEEVKFGGKVIDRGMGQSRV 62

Qy 61 DENTLDLIVTNAMIIDYTYGIKADIGKNGKIHGIGKAGNKMDQGYSPHMVVGVTAL 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 AADVPDVTIVNAVLDLHWGIVKADIAIKHGRIAAIGKAGNPDIQGVTV--IAIGAATEVI 120

Qy 121 AGEQMITAGGIDSHTHFLSPQFPPTALANGVTVMFGGTPGVDTGNATTITPGKWNLHR 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 AGEGLIVTAGGIDTHIHFISPOQIDREALSGVTVMFGGTPGATGNATTCTGPFWMHR 180

Qy 181 MLRAAEBSYMNVGFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 MLQAADGWPINLGLFGKGNASLPQLPVEQIAAGAIGLKLHEDWGTTPPAIDNCLSVAD 240

Qy 241 DVQVCIHTDTWNEAGYDDTLNANNGRAIHAYHIEGACGSHSPDVTWAGHENILIPSS 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 DTQVAIHTDLNEAGFYESTVAAPFKGTIHTYHTGAGGGGHAPDILKVCGE MNVLP 300

Qy 301 PTIPVTINTVAEHLDMLTCHHLDKRIRLEDLPQSQRIRPQSIAAEDVLHDMGVIA 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 PTRPYTINTLDEHLDMLMVCCHLDPSIAEDLAFAESIRRTIAAEDILHDLGALM 360

Qy 361 DSQAMGRAGEVPIRTWTQADKNKKEFGKLPDGDKNDNFRIKRYISKYITINPALTH 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 DSQAMGRGEVIIRTWQTAHKMKYQRGALPEDTARNONFRAKRYVAKYITINPAL 420

Qy 421 YIGSVEGKIADLVVNNAPFGVTKPKVIKGGWVVFSEMGDSNASVPTPQPVYREMF 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 EVGSIETGKADLVLEWPAFFGFKPISMLKGMALAQMGDPNASIPTPQPVHYREMF 480

```

RESULT 15
US-10-282-122A-47638
; Sequence 47638, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:44:57 ; Search time 34 Seconds

(without alignments)
1124.507 Million cell updates/sec

Title: US-09-904-994b-3

Perfect score: 2999

Sequence: 1 MKMKQEVNTYGTGDKV.....KLCTSKPTSQVPLAQRYYTF 568

Scoring table:

Gapop 4.0 , Gapext 1.0

Searched: 239914 seqs, 67312017 residues

Total number of hits satisfying chosen parameters: 239914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US09_NEW_PUB.pap:*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US06_NEW_PUB.pap:*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US07_NEW_PUB.pap:*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US08_NEW_PUB.pap:*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubpaa/PCT_NEW_PUB.pap:*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10_NEW_PUB.pap:*
- 7: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US11_NEW_PUB.pap:*
- 8: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1797	59.9	571	6	US-10-471-571A-3300
2	1794	59.8	653	6	US-10-953-349-33827
3	1794	59.8	784	6	US-10-953-349-33826
4	1794	59.8	841	6	US-10-953-349-33825
5	1769	59.0	843	6	US-10-449-902-47195
6	1748	58.3	836	7	US-11-292-431-37
7	316	10.5	2314	7	US-11-013-711-11
8	315	10.5	572	6	US-10-505-928-98
9	315	10.5	572	7	US-11-301-554-1815
10	309	10.3	1629	6	US-10-471-571A-3542
11	308	10.3	1006	6	US-10-449-902-41462
12	308	10.3	1705	7	US-11-304-590-3
13	305	10.2	7362	7	US-11-330-403-9773
14	303	10.1	801	6	US-10-471-571A-5160
15	303	10.1	4834	6	US-10-505-928-827
16	302	10.1	636	6	US-10-471-571A-1606
17	301	10.0	1349	6	US-10-471-571A-3352
18	301	10.0	2053	7	US-11-013-711-9
19	299	10.0	753	6	US-10-449-902-45171
20	298	9.9	498	6	US-10-953-349-4543
21	298	9.9	709	6	US-10-953-349-4542
22	298	9.9	725	6	US-10-953-349-4541
23	297	9.9	1992	7	US-11-013-711-3
24	297	9.9	1992	7	US-11-013-711-13
25	297	9.9	2047	7	US-11-013-711-4

26	297	9.9	2047	7	US-11-013-711-7	Sequence 7, Appli
27	295	9.8	951	6	US-10-449-902-41597	Sequence 41597, A
28	295	9.8	1247	7	US-11-330-403-18943	Sequence 18943, A
29	295	9.8	1272	7	US-11-330-403-3001	Sequence 3001, Ap
30	294	9.8	546	7	US-11-330-403-14342	Sequence 14342, A
31	294	9.8	1084	7	US-11-121-154-115	Sequence 115, App
32	293	9.8	928	7	US-11-330-403-10515	Sequence 10515, A
33	293	9.8	2499	7	US-11-070-573-26	Sequence 26, Appl
34	292	9.7	574	7	US-11-330-403-17416	Sequence 17416, A
35	292	9.7	1531	7	US-11-333-747A-44	Sequence 44, Appl
36	291	9.7	539	6	US-10-449-902-50636	Sequence 50636, A
37	291	9.7	574	7	US-11-330-403-9966	Sequence 9966, Ap
38	291	9.7	987	6	US-10-449-902-45475	Sequence 45475, A
39	291	9.7	1167	7	US-11-121-154-1	Sequence 1, Appli
40	290	9.7	563	7	US-11-056-355B-83788	Sequence 83788, A
41	290	9.7	565	7	US-11-056-355B-83787	Sequence 83787, A
42	290	9.7	566	7	US-11-056-355B-83786	Sequence 83786, A
43	290	9.7	2740	7	US-11-330-403-5490	Sequence 5490, Ap
44	290	9.7	3016	7	US-11-330-403-7229	Sequence 7229, Ap
45	290	9.7	22152	6	US-10-544-944-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-471-571A-3300
; Sequence 3300, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:

; APPLICANT: CHIRON Spa

; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS

; FILE REFERENCE: P026927W0

; CURRENT APPLICATION NUMBER: US/10/471,571A

; CURRENT FILING DATE: 2003-09-12

; PRIOR APPLICATION NUMBER: GB-0107661.1

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 5642

; SOFTWARE: SeqWin99, version 1.03

; SEQ ID NO 3300

; LENGTH: 571

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1)..(571)

; OTHER INFORMATION: urease alpha subunit

US-10-471-571A-3300

Qy	2	KMKQEVNTYGTGDKVRLGDTDLMAVEHDYTYGHELFKFGAGKTIREGQGQSNP- 60	Query Match	59.9%	Score 1797;	DB 6;	Length 571;
Db	4	KMTQTSYLSYGTGDSIRLGDITLFAQIEKDYAVYGEATFGGKSIKRDGMAQ-N-PR 61	Best Local Similarity	57.5%	Pred. No. 1.1e-35;		
Qy	61	-----DENTLDLIVTAMLIIDYTGIVKADIGIKNGKHIGIKAGNKMDQGVSPHVVGVG 116	Matches	330;	Conservative	97;	Mismatches 134;
Db	62	VTRDDVNVADIVISNAVIIIDYDKVKADIGIKNGYIFAIGNAGNPDIMDND--IIIGST 119				Indels	13;
Qy	117	TEALAGEMIITAGIDISHTHFLSPQPPPTALANGVTTMFGGGTGPVDCGTNATTITPPGW 176				Gaps	6;
Db	120	TDITAAEGKIVTAGIDITHVHFINPEQAEVALESITTHIGGTCASEGSKATVTPGW 179					
Qy	177	NLHRMLRAABEYSNMVGLFGKNSSSKKQLVQVQVBAIGFKLHEDWGTTPSAIDHCLSV 236					
Db	180	HIHRMLEAELGPIINVGTGKQATNPATLIEQINAGAIGLKVHEDWGATPSALSHALDV 239					
Qy	237	ADEYDVQVCHTDTVNEAGYVDDTLNANNGRAIHAYHLEGAGGHSPOVITWAGELNIIPL 296					
Db	240	ADEFDVQIALHDLTNEAGFMEDTMAAVKDRVLHMYHTEGAGGGHAPDLIKSAAFSNILP 299					

Db 756 EYTVTADGEVLTCQAPTLPLSRNYFLF 784

RESULT 4

US-10-953-349-33825

; Sequence 33825, Application US/10953349

; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; FILE REFERENCE: ENCODED THERY

; CURRENT APPLICATION NUMBER: US/10/953,349

; PRIOR FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: Patent in version 3.3

; SEQ ID NO 33825

; LENGTH: 841

; TYPE: PRT

; ORGANISM: Zea mays subsp. mays

US-10-953-349-33825

Query Match 59.8%; Score 1794; DB 6; Length 841;

Best Local Similarity 58.9%; Pred. No. 3.3e-35;

Matches 335; Conservative 89; Mismatches 139; Indels 6; Gaps 2;

Qy 6 QEYVNTYPTKGDVKRLGDTDLMAEVEHDYTYTYGEEKFKGAGKTIREGMQO-SNSPDENT 64

Db 273 EKVASIYPTTGDGKIRLGDTDLMAEIEKDFAPYGDCEIFGGKVLRDGMQASGYPESFC 332

Qy 65 LDVITNAMIIDYTYGIKADIGIKNGKHGKAGKDMQGVSPHVMVGVGTALAGEG 124

Db 333 LDTVTNNAVVIDYTYGIKADIGIKGGLIIVAGKAGNPDVMDGVNMMIVGVNTEVIASEG 392

Qy 125 MITTAGGIDSHHTPLSPQQPTALANGVTTFMFGGTGPDGTGATTTTPGKNNLHRLRA 184

Db 393 MITTAGGIDCHVHFICPLQAEAEIATSGITTLVGGTGPAGHGTCACTTPASQKLMLQS 452

Qy 185 ABEYSNMVGLKGNSSSKQLVQVEAGATGFKLHEDWGTPSAIDHCLSVADSDYDVQV 244

Db 453 TDQLPINWGTGKNTSKPEGLAIIKAGAMGLKHLHEDWGTPSAIDNCLSVADSDYDVQV 512

Qy 245 CIHTDTVNEAGYDDTLNMMNGRAIHAYHIEGAGGHSPPDVTWAGELNIIPLSSTTPTIP 304

Db 513 NIHTDTLNEGCVHEHTIAAFKRAIHTYHSEGAGGCHAPDIKVCVGNVLPSSNTPTRP 572

Qy 305 YTIINTVAEHLDMLTCHLDKRIREDLQFSQIRPGSIAAEDVLHDMGVITAMTSSDQA 364

Db 573 FTSNTVDEHLDMLMVCCHLDKNIPEDEVAFABSRIRAEITAAEDILHDMGAISIISSDQA 632

Qy 365 MGRAGEVTPRTWOTADKNKKEFGKLPEDG-----KNDNFRIKRYISKYITINPALTHGVS 419

Db 633 MGRVGEVITRTWQANKVQSGSLPSSGDANAPDSNLRIRIYIAKYITINPAIVNGFS 692

Qy 420 EYIGVSEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFG 479

Db 693 DFVGSVEVGLADLVLPKPSFGAKPELVKVGGAIAWANMGDPNASIPTPEPVVVRPMFG 752

Qy 480 HHGAKAFDTSITFVSKVAYENGVEKGLERQVLPVNCNRNITKKDKFNDKTKAKITVDP 539

Db 753 AFGKAGSSNSIAFVSKAAKEAGVATEYRLKREAVGRVRLGLTKLMDKLDALPKIEVDP 812

Qy 540 KTFEVEFDGKLTCKTSKPTSOVPLAQRITFF 568

Db 813 EYTVTADGEVLTCQAPTLPLSRNYFLF 841

RESULT 5

US-10-449-902-47195

; Sequence 47195, Application US/10449902

; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.

; APPLICANT: Bio-oriented Technology Research Advancement Institution.

; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205YI-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; PRIOR FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 47195

; LENGTH: 843

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-449-902-47195

Query Match 59.0%; Score 1769; DB 6; Length 843;

Best Local Similarity 58.1%; Pred. No. 1.3e-34;

Matches 333; Conservative 91; Mismatches 137; Indels 12; Gaps 7;

Qy 6 QEYVNTYPTKGDVKRLGDTDLMAEVEHDYTYTYGEEKFKGAGKTIREGMQO-SNSPDENT 64

Db 273 EKIASMYGTTGDKIRLGDTDLFAEIEKDYAIYGDCEIFGGKVLRDGMQASGYPASDC 332

Qy 65 LDVITNAMIIDYTYGIKADIGIKNGKHGKAGKND-M-QDGVSPHVMVGVGTALAGEG 122

Db 333 LDTVTNNAVVIDYTYGIKADIGIKGGLIIVAGKAGNPDVMDGVNMMIVGVNTEVIAA 392

Qy 123 EGMITAGIDSHHTPLSPQQPTALANGVTTFMFGGTGPDGTGATTTTPGKNNLHRLM 182

Db 393 EGMITAGIDCHVHFICPLQAEAEIATSGITTLVGGTGPAGHGTCACTTPSPSHMKML 452

Qy 183 RAAEYSNMVGLKGNSSSKQLVQVEAGATGFKLHEDWGTPSAIDHCLSVADSDYDV 242

Db 453 QSTDLPINWGTGKNTTKPDGLAIIKAGAMGLKHLHEDWGTPSAIDNCLSVADSDYDV 512

Qy 243 QVCIHDTVNEAGYDDTLNMMNGRAIHAYHIEGAGGHSPPDVTWAGELNIIPLSSTTPT 302

Db 513 QVNIHTDTLNEGCVHEHTIAAFKRTIHTYHSEGAGGCHAPDIKVCVGNVLPSSNTPT 572

Qy 303 IPTYINTVAEHLDMLTCHLDKRIREDLQFSQIRPGSIAAEDVLHDMGVITAMTSSDS 362

Db 573 RPTLTNTVDEHLDMLMVCCHLDKNIPEDEVAFABSRIRAEITAAEDILHDMGAISIISSDS 632

Qy 363 QAMGRAGEVTPRTWOTADKNKKEFGKLP-----ED-KNDNFRIKRYISKYITINPALTHG 417

Db 633 QAMGRAGEVTPRTWOTADKNKKEFGKLP-----ED-KNDNFRIKRYISKYITINPAIVNG 692

Qy 418 VSEYIGVSEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREM 477

Db 693 FSDPFGVSEVGLADLVLPKPSFGAKPELVKVGGAIAWANMGDPNASIPTPEPVVVRPM 752

Qy 478 FGHGAKAFDTSITFVSKVAYENG--VKEKGLERQVLPVNCNRNITKKDKFNDKTKAKI 535

Db 753 FGAFCGAGSANSIAFVSKAAKEAGVAVOYKLG--KREAVGRVRLGLTKLNMKLDALPKI 810

Qy 536 TVDPKTFEVDGKLTCKTSKPTSOVPLAQRITFF 568

Db 811 DVDPEYTVTADGEVLRCQPTPTPLSRNYFLF 843

RESULT 6

US-11-292-431-37

; Sequence 37, Application US/11292431

; Publication No. US20060121061A1

; GENERAL INFORMATION:

; APPLICANT: COLE, Garry T.

; APPLICANT: CHEN, Xia

; APPLICANT: SESHAN, Kalpathi R.

; APPLICANT: HUNG, Chiung-Yu

```

; APPLICANT: XUE, Jiamin
; APPLICANT: YU, Jieh-Juen
; TITLE OF INVENTION: Attenuated Vaccine Useful for
; TITLE OF INVENTION: Immunizations Against Coccidioides spp. Infections
; FILE REFERENCE: 52952200600
; CURRENT APPLICATION NUMBER: US/11/292,431
; CURRENT FILING DATE: 2005-12-02
; PRIOR APPLICATION NUMBER: US 60/633,399
; PRIOR FILING DATE: 2004-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Coccidioides posadasii
US-11-292-431-37

Query Match      58.3%; Score 1748; DB 7; Length 836;
Best Local Similarity 57.2%; Pred. No. 4.2e-34;
Matches 327; Conservative 93; Mismatches 136; Indels 16; Gaps 10;

Qy      3  MKQEVYNTYPTKGDVRLGDTDLWAEVHDYTYGEEKFGAGKTIREGMGQS---NS 59
Db      272 MTRAYARMFGTTGDVVKGDTDLWIKVEKDLTYGDECFGGKTIRGGMGQATGRHS 331

Qy      60 PDENTLDLVITNAMIIDYTIYKADIGIKNGKIHGKAGKMDQGVSPHVMVGVGTEA 119
Db      332 VD--VLDTVLVNALLVDWYGIYKADIGLKGDLGKAGNPDMDGVTNNIVGSSTDV 389

Qy      120 LAGSMIITAGDISHTFLSPQPPPTALANGVTMTFGGGTGPVDGVTNATTIPGKNLH 179
Db      390 IACGKIIVTAGIDITHVHFICPQVVEALASGVTLLGGGTGPTGEGSNATTCTPAP-NQF 448

Qy      180 R-MLRAAEYSNMGFLGKNSKKQLVEQVAGAIKFLHEDWGTTPSALDHCLSVAD 238
Db      449 KTMQACDHLPLNVLGTGKNDGSLPLSDQCRAGAGLKVHEDWGATPAVIDTCLQVCD 508

Qy      239 EYDVQVC-IHTDTVNEAGYVDLTNMGRAIHAYHIEGAGGHSPPDVTIMAGELNIPLS 297
Db      509 EFDIQ-CLIHDTLINESFVQGTNAFNVRVHTYHTEGAGGHAPDIIISVVEKENVLP 567

Qy      298 STTPTPIYNTVAHLDMLTCHHLDKRIRIEDLQFSQSRIRPGSIAAEDVLHDMGVIAM 357
Db      568 STNTPRYVTNTLDEHLDWVCHHLSKDIPEDEVAFBSRISRTIAAEDVLHDTGALSM 627

Qy      358 TSSDSQAMGRAGEVTPRTWQADNKKKFKLPED-GKNDNFRKRYISKYITINPALTH 416
Db      628 LSSDSQAMGRGCEVVVVTWNTAHKNMERGRLEKDEGTDSDNFRKRYISKYITINPAIAQ 687

Qy      417 GVSEYIGVBEKGKADLVVWNPAPFGVPKIVIKGMVVFSEMGDSNASVPTPQPVYRE 476
Db      688 GMAHTIGSVFVGKTDLVLFKPFANFTSPVWLVKSGMAVSQMGDPNGSIPTIEPIWRP 747

Qy      477 MFGH-HGKAKFDTSTIFYSKVAYENGVEKGLGLERQVLVKNCRNITKDKPFNDKTAKI 535
Db      748 MYASLNPKA----SIMFVSQASIKGLIIDSYHLKRIEIPVKNCRNISKRDMMKFNIMPKN 803

Qy      536 TVDPKTFEYVFGKLTCKTSKPTSQVPLAQRYTF 567
Db      804 RVDPEYVVEADGEBCTEAPVSEPLTQDY-F 834

RESULT 7
US-11-013-711-11
; Sequence 11, Application US/11013711
; Publication No. US20060159709A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN OF MORAXELLA

```

```

; FILE REFERENCE: 1038-921MIS:Jb
; CURRENT APPLICATION NUMBER: US/11/013,711
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2314
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-711-11

Query Match      10.5%; Score 316; DB 7; Length 2314;
Best Local Similarity 25.7%; Pred. No. 2.8e+02;
Matches 196; Conservative 83; Mismatches 215; Indels 270; Gaps 90;

Qy      1  MKMKQEVYNTYPTKGD-KV-RLGD-TDLWAEVHDYTYGEEKL-K-----FGA--GKT 49
Db      587 VKITKDSGINA-----GDKISNVKDATD-----DTDAVTV-KOLKQVQDADGALQSFS 635

Qy      50 IREGMGQ-----SNSPDENTLDLVIT-----NAMI--DVT-GIYKADIGIK--NG-- 90
Db      636 IREKGGQFTISNLSNGNTPNTPF-TIIFAGENGISINDIAG--KVKVGIDPINGLET 692

Qy      91 --KIH-GI---GK-----AGN--KDMQDGVSPHM--VVGVG-----TEALAGEGMII 127
Db      693 TPKLTVGSDKDGKTKLVIEQVAGSNDTKNIIRGLSPTLPSITNAGGVRTTE---QNTI 748

Qy      128 T-----AGGI-D-SHTHF-LSPQFPPTALANGVTMT-FGGGTGVPDGTNATT--ITPG 174
Db      749 TSDEKSKAAASIGDILNTGFLNKNNSVGFVSTYNTVDF-----IDG-NATTAKVTYD 801

Qy      175 KNLHRLMRAAE-EYSMN-----GFLGKNSSSKKQLVEQVAGAIKFLHEDWGT 225
Db      802 ETN-----QTSKVTYDVNVDEKTIETLTDGNGKTKIGVK-TTTLTTTANG-KA-TNFST 853

Qy      226 TPS-AIDHCLSVADYDVQVC-IHT-----DT-----VNEAGYVDD-TL-----NAMNG 266
Db      854 TDNDALVNAKDIKENLNTLAKIHTTKGTADTALQTFKVKDGTATDEITVKGDKGTQNG 913

Qy      267 RAIHAYHIEGAGG---GHSPD-VITM-----AGE---LN-----IL-P-----296
Db      914 KTVNTLKLKGENGLFVATNKGDTVTFGINTQSLKAGDSTTLNKDGLSINKNPASNEQIQV 973

Qy      297 -----SST-----TPTIPTYT---IN-TVAE-HLDMLTCHHL--DK-RIED 330
Db      974 GADGVKFAKVDKNSSTGIDGTSRI--TKDQIGFTGANGSLD--TTKPHLTDKLKVGE- 1028

Qy      331 LQFSQSRIRPGSIAAEDVLHDMGVIAMTSSDSQAMGRA-----GEV---IPRTWQADNKN 383
Db      1029 VEITWYGINAGKKITNI--QSGDITQNSNDATVGRVVDLKTLELSKNSAKTAQNSL 1086

Qy      384 KEFGKLPEDGKNDNFRIKR-YISKYITINPALTHGVSEYIGVSEEGKIADLVVWNPAP-- 440
Db      1087 HEFSVADEQG--N-HFTVSNPY-SSY--DTSKTSVDITFAG--ENG-ITTKV--NKGVVR 1135

Qy      441 FCV---K-----PKIVI---KG-GMVVFSEMGD-----SN--ASVPTPQPVYREMGH- 480
Db      1136 VGIDQTKGLTTPKLTGVGNNGKGIVIDSKDGQNTITGLSNTLANV-T-----NDGAGHA 1188

Qy      481 --HGKAKFDT-----SITFVSKVAVE---NGVEKGLGLERQVLVKNCRNITKDKF-KF 528
Db      1189 LSQGLAN-DTDKTRAASIGDVLNAGFNQNG--EAVDF-----VSTDVTV---DFIDG 1236

Qy      529 NDKTAKITVD--PKT-----FEVFDGKLTCKTSKPTSQVPLAQRYT 566
Db      1237 NATTAKVYDTSKTSKVYDVNDNK--TIEVTSKGLGVKTT 1278

RESULT 8
US-10-505-928-98
; Sequence 98, Application US/10505928

```

Publication No. US20060088532A1

GENERAL INFORMATION:

APPLICANT: Ludwig Institute for Cancer Research et al.

TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES

FILE REFERENCE: 28967/39178

CURRENT APPLICATION NUMBER: US/10/505,928

CURRENT FILING DATE: 2004-08-27

PRIOR APPLICATION NUMBER: US 60/363,019

PRIOR FILING DATE: 2002-03-07

NUMBER OF SEQ ID NOS: 866

SOFTWARE: PatentIn 3.2

SEQ ID NO 98

LENGTH: 572

TYPE: PRT

ORGANISM: Homo sapiens

US-10-505-928-98

Query Match 10.5%; Score 315; DB 6; Length 572;
Best Local Similarity 23.4%; Pred. No. 9.6; Indels 185; Gaps 61;
Matches 147; Conservative 94; Mismatches 201;

QY 55 GQSNPDLTLD-LVITNAMI-DYTGKADIGIKNGKIHGIGKAGNKMDDGVSPHMV 112
DB 5 GKKSIP-HITSRLIKGRIINDQSLY-ADVLEDGLIKQIGE-N-----LI 50
QY 113 V-GVTEALAGEMIITAGGIDSHTHFLSPQQFPPTALAN-----GV-TTMFGGTPVD-- 164
DB 51 VPG-GVKTEANGRMVPGGIDVNTYLOKPSQGMTA-ADDFQGTAAALVGGTTMIIDHV 108
QY 165 ----GTTNATTITPGKWNHRLMLRAE-----EYSNVGFLG--KGNSSKKQLVEQVEAG 213
DB 109 VPEGSSLLT-SFEKW--H---EAADTKSCCDYSLHVDITSWYDGVRELEVLVQ--DKG 160
QY 214 AIGFKL--HED-WGTTPTSAIDHCLSVADVDVQVCIHTDTVNEAGYVDDTLNAMNGRAI 269
DB 161 VNSFQVWAYKDVQMSDSQLYEATFLKGLGAVILVHAE--N-G---D-LIAQEQKRI 212
QY 270 HAYHIE-GAGG--GHS-----PD-----VITMAGELN-----ILPSSSTTPTI--- 303
DB 213 ----LEMGITGPEGHLSRPELEAEAVFRAITTIAGRINCPVITKVMKSAADIILAR 268
QY 304 ---PYTI-NTVAEHLMDLMTCHLDKRIREDLQFSQS--RIRPGSIAAEDVLHDMGVAMT 358
DB 269 KGGLVFGEPITAAASLGTDGT-HYNSKNWAKAAAFVTPPLSPDP-TTPDYLTSL-----LA 322
QY 359 SSSQAMGRAGEVTPRTWQADK--NKKEGKLEP-----D-----GKONDN 398
DB 323 CGDLQVTS-SGH-CP--YSTAQKAVGKDNFTLIEGVNGIEERTVTVWDKAVATGRMDEN 378
QY 399 FRIKRYISKYTINPALTGHVSEYIGSVBEGKIADLVVWNP-----A-----F 440
DB 379 ----QFVAVTSTNAKIFNLVPRKGRITAVGSDADVIVDPDKLTIITAKSHKSAVEYNIF 434
QY 441 FGK--PKIVKGMVVFSEMGDSNAS-----VPT---PQPVYR-----EMFGHHG 482
DB 435 EGMECHSPLVVISQKIVF-EDGINVNGKMGRFIPKAPPEHLQVQKIRNKVFLQ 493
QY 483 --KAKPDTSTFV-SKAYEN---GVKEGLERQVLPVKNCRNITKDKPFKNDKAKIT 536
DB 494 VSRGMYDGPVYEPATPKYATPAPSASKSPS-KHQPPI---RNLHQSNSLSG--AQID 547
QY 537 VD--P-KTEFV--DGKLCSTKPTS 557
DB 548 -DNNPRKTHRIVAPGGR---SNITS 570

RESULT 9

US-11-301-554-1815

Sequence 1815, Application US/11301554

Publication No. US20060088527A1

GENERAL INFORMATION:

APPLICANT: Henderson, Robert A.

APPLICANT: Wang, Tongtong

APPLICANT: Watanabe, Yoshihiro
APPLICANT: Kaloe, Michael D.
APPLICANT: Sleath, Paul R.
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Durham, Margarita
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: McNabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C21
CURRENT APPLICATION NUMBER: US/11/301,554
CURRENT FILING DATE: 2005-12-13
PRIOR APPLICATION NUMBER: US 10/283,017
PRIOR FILING DATE: 2002-10-28
PRIOR APPLICATION NUMBER: US 10/113,872
PRIOR FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 10/017,754
PRIOR FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: US 09/902,941
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 09/849,626
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: US 09/736,457
PRIOR FILING DATE: 2000-12-13
PRIOR APPLICATION NUMBER: US 09/702,705
PRIOR FILING DATE: 2000-10-30
PRIOR APPLICATION NUMBER: US 09/677,419
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 09/671,325
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US 09/658,824
PRIOR FILING DATE: 2000-09-08

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 2157

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1815

LENGTH: 572

TYPE: PRT

ORGANISM: Homo sapiens

US-11-301-554-1815

Query Match 10.5%; Score 315; DB 7; Length 572;
Best Local Similarity 23.4%; Pred. No. 9.6; Indels 185; Gaps 61;
Matches 147; Conservative 94; Mismatches 201;

QY 55 GQSNPDLTLD-LVITNAMI-DYTGKADIGIKNGKIHGIGKAGNKMDDGVSPHMV 112
DB 5 GKKSIP-HITSRLIKGRIINDQSLY-ADVLEDGLIKQIGE-N-----LI 50
QY 113 V-GVTEALAGEMIITAGGIDSHTHFLSPQQFPPTALAN-----GV-TTMFGGTPVD-- 164
DB 51 VPG-GVKTEANGRMVPGGIDVNTYLOKPSQGMTA-ADDFQGTAAALVGGTTMIIDHV 108
QY 165 ----GTTNATTITPGKWNHRLMLRAE-----EYSNVGFLG--KGNSSKKQLVEQVEAG 213
DB 109 VPEGSSLLT-SFEKW--H---EAADTKSCCDYSLHVDITSWYDGVRELEVLVQ--DKG 160
QY 214 AIGFKL--HED-WGTTPTSAIDHCLSVADVDVQVCIHTDTVNEAGYVDDTLNAMNGRAI 269
DB 161 VNSFQVWAYKDVQMSDSQLYEATFLKGLGAVILVHAE--N-G---D-LIAQEQKRI 212
QY 270 HAYHIE-GAGG--GHS-----PD-----VITMAGELN-----ILPSSSTTPTI--- 303
DB 213 ----LEMGITGPEGHLSRPELEAEAVFRAITTIAGRINCPVITKVMKSAADIILAR 268
QY 304 ---PYTI-NTVAEHLMDLMTCHLDKRIREDLQFSQS--RIRPGSIAAEDVLHDMGVAMT 358
DB 269 KGGLVFGEPITAAASLGTDGT-HYNSKNWAKAAAFVTPPLSPDP-TTPDYLTSL-----LA 322

QY 359 SSDSQMGRAGEVPIRTWOTADK--NKKEFGKLP-----D-----GKNDN 398
 Db 323 CGDLQVTS--SGH-CP--YSTAQKAVGKDNFTLIEGVNGIERMTVVDKAVATGKMDN 378
 QY 399 FRIKRYISKYITINPALTGHVSEYIGSVBEGKIADLVVNP-----A-----F 440
 Db 379 ----QFVAVSTNAAKFNLYPRGRIAVGSDADVVDKLTITIAKSHKSAVEYNIP 434
 QY 441 FGVK-----PKIVIKGMVVFSEMGDSNAS-----VPT---PQPVYR-----EMFGHHG 482
 Db 435 EGMECHGSPLVVISQKIVF--EDGNINVNKGWGRFIPKAPFELHYQKVRKRNKVFGIQQ 493
 QY 483 --KAKPDTSTIV--SKVAYEN--GVKEKLGLEQVLPVKNCRNITKDKFNDKTAKIT 536
 Db 494 VSRMGDPVVEVPATPKYATPAPSASSPS--KHQPPPI---RNLHOSNFSLSG--AQID 547
 QY 537 VD--P-KTFEVFV--DGKLCSTKPTS 557
 Db 548 -DNNPRITGHRIVAPPGR---SNITS 570

RESULT 10

US-10-471-571A-3542
 ; Sequence 3542, Application US/10471571A
 ; Publication No. US20060115490A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SA
 ; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE: P026927WO
 ; CURRENT APPLICATION NUMBER: US/10/471,571A
 ; CURRENT FILING DATE: 2003-09-12
 ; PRIOR APPLICATION NUMBER: GB-0107661.1
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 5642
 ; SOFTWARE: SeqWin99, version 1.03
 ; SEQ ID NO 3542
 ; LENGTH: 1629
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 ; FEATURE:
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (1)..(1629)
 ; OTHER INFORMATION: hypothetical protein
 US-10-471-571A-3542

Query Match 10.3%; Score 309; DB 6; Length 1629;
 Best Local Similarity 25.6%; Pred. No. 1.7e+02;
 Matches 161; Conservative 80; Mismatches 199; Indels 190; Gaps 74;
 QY 5 KQEVV-N--T--YGPT-KGDKV-RLGDT-----DLWAEVE-HDYTTYGEELKFGAGKTIR 51
 Db 90 KSEVTSNKDTIEHPSPVKAEDISKEDTPKEADV-AEVQPKSSVTHNAE---TPK-VR 143
 QY 52 EGMQSQSPDENTLDLVITNAMIIDYGIYKADIGIKNGKHIGKAGNKMQD---D-G 106
 Db 144 ----KARSVDGSEFDITRDSKNVVEST-----PITIQ-GKEHFEG-YGSVDIQKPTDLG 192
 QY 107 VSPHVVGVGTEALAGEGMILTAGID--SHTHFLSPQFPPTALANGVTTMFGGSGTPVD 164
 Db 193 VSEVTRFVNGNES---NGLI---GALQJKNKIDFSKOPNFKVRVANNHQS---NTTG-AD 242
 QY 165 GTNATTITPGKWNLRMLRA-AEEYSMVVGLG-KGNSSSKQLVQVEAGAGFKLHED 222
 Db 243 G-----WGF-LPSKGNAAEVLITNGILGDKG-----L---VNSG-GFKI--D 277
 QY 223 WGTTPSAIDHCLSVADYDVQVCIHTDFTVNBAGYVDDT-LNAMNG-RAIHAYHIEGAGG 280
 Db 278 TG-----Y-----IYTS-----MDKTEKQAGQYRGYGF-VKNDSSG 310
 QY 281 HSPDVITMAGELNIPSSTTTPIPTINTVAEHLDMLTCHLDKXIRED--LQF--SQS 336
 Db 311 NS-----QMVGE-NI--DKSKTNFLNADNS-TNTSDGKF---H-GQRL-NDVILTYVASTG 358.

QY 337 RIRPSIAA---EDVLHDMGV-----IAMTSSDSQAMG-RAGEVPIRTWOTADKNKKE 385
 Db 359 KMR-ABYAGKTWETSITDLGLSKNQAYNFLITS--SORWGLNQG-INANGMMRTDLKGE 414
 QY 386 FGKLP-----DGK-DNDNFRIKRYISKYITINPALTGHVSEYIGSVBEGKIADLVV 436
 Db 415 FTFTPEAKTITTELEKKVEEIPFKER---KF--NPDLAG-TEKV--TREGQKGEKIT 466
 QY 437 NPAPFVGK-PKIVIKGMVVFSEMGDSNASVPTPOV-----YYREMF--GHHGKAKFDT 488
 Db 467 TPT---LKNP---LTKG--VIISK-GEPKKEI-TKDPINELTEYGPETIAPGH--RDEFPD 514
 QY 489 SITFVSK--VAYENGVE-KLG-LER-----QVL---PVQNCRNITKDKFNDKTAKITV 537
 Db 515 KLPTGEKEVPGKIPKIPETGTVVRPPVDSVTKYGPVKGDSIVEKEEIPF-EKERKFN 573
 QY 538 D--PKTFEVFVDG---KLCTSKPTSOVPL 561
 Db 574 DLAPGTEKVTREGQKGEKITIT--PTLKNP 602

RESULT 11
 US-10-449-902-41462
 ; Sequence 41462, Application US/10449902
 ; Publication No. US20060123505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: National Institute of Agrobiological Sciences.
 ; APPLICANT: Bio-Oriented Technology Research Advancement Institution.
 ; APPLICANT: The Institute of Physical and Chemical Research.
 ; APPLICANT: Foundation for Advancement of International Science.
 ; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
 ; FILE REFERENCE: MOA-A0205Y1-US
 ; CURRENT APPLICATION NUMBER: US/10/449,902
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: JP 2002-203269
 ; PRIOR FILING DATE: 2002-05-30
 ; PRIOR APPLICATION NUMBER: JP 2002-383870
 ; PRIOR FILING DATE: 2002-12-11
 ; NUMBER OF SEQ ID NOS: 56791
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 41462
 ; LENGTH: 1006
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-10-449-902-41462

Query Match 10.3%; Score 308; DB 6; Length 1006;
 Best Local Similarity 25.5%; Pred. No. 57;
 Matches 167; Conservative 89; Mismatches 187; Indels 211; Gaps 77;
 QY 10 NTY-GPTKGDKVR--LGDTDLWAE-VEHDYTTYGEELKFGAGK-TIR-----E---GMG-Q 56
 Db 74 DTFLGAGADTWQGLGD-DAFTDNMIGDYT--GDVIDGLNDTVRFFLSERATGVAF 130
 QY 57 SNSPDENTLDLV---ITN--AMIID--YTGIVKADIGIKNGK-----IHGIGKAGNKM 105
 Db 131 AADPTE-ITVTLFGAKLHNEIYQVVDGNTG---AD-NMTGGRWNDFQNGGADRLDGG 185
 QY 106 GYSPHVVGVGTEAL-AGEGMILTAGIDSHTHF-----LSPQFPPTAL--ANGVTM 155
 Db 186 G-ADYLSGGAGADVLTIGGGQDYINGGDDKDLFGYGGDNLSSGGNGDDTIEGAGDAL 244
 QY 156 FG--G-----GTGPVD-----GTNATTITPGKWNLRMLRAAEEYSM---NVGFL-KG 198
 Db 245 YGEDGNDKLAVAGAGN-DYLSGGAGSDTLGG-----AGDYLITGDNKVLGSG 294
 QY 199 N-----SSSKQL-----VEQVEA-----GAIGFKLH-----EDWGTTPSAIDHCLSV 236
 Db 295 NDQILATVNSGKSIDGGSGVDRLELRVYNGAFTAKAENLVNTLSD-GTT---IQH---- 346
 QY 237 ADEYDVQVCIH-----TDTV-----NEAGY-VDDTLNAMNGR-AIHAYHIEGAG----- 278
 Db 347 VEAYDV-----HGSNASNQFT-TLGGDDNLGYGGDDTLRAGAGNDSIFA-----GDGRDKL 397

Db 5282 DIIKWTSTSEDIFGDNAYILRAQNEASSDQWMLQAKTLVNDKG-GDDTIITSGG 5340
Qy 267 RAI-----HAYHIEGAGGHSPPVITMAGEL-----N-IL--PSSTPTTI--PYTINTV 310
Db 5341 KVLVGGGADGSAINGDHL--VIGDNGQLDYKDSENGILRKMQSLDNVIGGEDTI-TL 5397
Qy 311 AEHLDMLTCHHLKRIREDLOFQSRRPQSIARAEVDLHDM-GVIAMTS-SDSQAMGRA 368
Db 5398 REGFVVMGGY--DDEI---TVYTE-----TVAASGVKDLTGIVSVASVSELEAKGRT 5447
Qy 369 GEVTPRTWQTADKNKCF---G-KL-PED-----GKNDNFRKRIKVIKTYTIN---P 412
Db 5448 GRYI-----TGDNMEVDPLHGLYKIOSNDAVAMGGD-DTISG---NRATTNDLGYQ 5498
Qy 413 ALTHGV-SEYIGSVBEGKIA-DLV-----VW-----NPAPF--GVKPKIVIKGMMVFSBMG 460
Db 5499 VLIQMAADTI-TVEDTSTQDVFGDNVYVERSTSSYFHMINDSTLNQGG-----G 5550
Qy 461 DSNASVPTQPYYREMPHHGKAKPDT-----SITFVKVAYENGKELGLERQVLVVK 516
Db 5551 D---TIVTGOGEKW--IVGGMG-A--DTAEVVSINLTDTLFA-GDSVDITFSR-----AK 5597
Qy 517 --NCRNITKDKFKE--NDKTAKITVDPKTFEV-FVDGK---L---CTSKPTSQV----- 559
Db 5598 TGNLTSTSTGLYFGGND--ALVVGDG---DVTFTGGKNDNSLIVDSTGATTAEVLAGTA 5652
Qy 560 --PLAQRYTF 567
Db 5653 GDDPNQRFAP 5662

RESULT 14

US-10-471-571A-5160
; Sequence 5160, Application US/10471571A
; Publication No. US20060115490A1

GENERAL INFORMATION:

; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0

; CURRENT APPLICATION NUMBER: US/10/471, 571A

; CURRENT FILING DATE: 2003-09-12

; PRIOR APPLICATION NUMBER: GB-0107661.1

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 5642

; SOFTWARE: SeqWin99, version 1.03

; SEQ ID NO 5160

; LENGTH: 801

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1)..(801)

; OTHER INFORMATION: nitrite reductase

US-10-471-571A-5160

Query Match 10.1%; Score 303; DB 6; Length 801;

Best Local Similarity 25.5%; Pred. No. 43;

Matches 171; Conservative 83; Mismatches 172; Indels 244; Gaps 84;

Qy 9 VNTYGPTKGDKVRIGDGLDPLA--E-VHEDYTYTG--EBLKFGAGK-----TIREGMGQS 57

Db 255 VNR-GIVVNDVMQTSNVAVGECAEHNGKVGLVAPL-YEQGKVLADHLTKETNGYK 312

Qy 58 NSPDNTL-----DL-----VITNMI--ID-YTG-----YKADIGIKNGKIHGKA--G 99

Db 313 GSTTFTSLKVGCDLISAGQIVENAEIKGIEIFNSVDNNYK-KIFLKGDNV--VGVAVLYG 369

Qy 100 NKDMODGVSP--HMV-VGVGTE--ALAGEGMIIITAGGIDSHTHFLSPQGFPTA-LANGVT 153

Db 370 --DIDDD--SRFNNMKKGESTEDYTLVS---LLTKGGEAS---LS-----IADWADDET 415

Qy 154 TWFGGGTGPVD-GT--NATTITP-GKNLHRLRAAEYSNMVFLGK-GNSSSK-KQIV 207

Db 416 IC--CGNG-VDKGTIVNA--ITENG-----FTTVEEVTAKT-----KAGNSCKCKPQI 459
Qy 208 EQVEAGAIGFKLHEDW-GTTPSAIDHCLSA-DEYDVQVCIIHTDTTNEAGYVDDTLNAMN 265
Db 460 AQILQHTLIG---DDFVAAKPAIGCGCTDLTRDQ-----IVTQ-IRAKG-----LK--T 501
Qy 266 GRAI-HAYHIEGAGGHSPPVITMAGELNLPSSSTPTTPTTINTVAEHLDMLTCHHL 324
Db 502 SKEVRHVLNFKNKG-----CP-KCRPAINYILNMVYPH-D-----HED 538
Qy 325 KR-----IREDLQFS---QSR-----IRPGSIAAEDVLHDMGVAMTSS 360
Db 539 ERESFANERHANLQNDGTFFVIPQMRGCVTDADQLIRLGEVAKK--YH-VPLVKVTG- 594
Qy 361 DSQAMGRAG---EVIPRTWQTAD---KNKKEFKLKPED-----GKNDNPF----- 399
Db 595 -SORVGLYGVKKEELPNWE--DLGMRASASAYGKTKRSVKSCVGFCEFRGTQVTTRLGI 651
Qy 400 RIKRIYIKTYTINPALTH-----GVSEYIG--S-VERGKIAD--LV-VNPNAPFVGPVKIV 448
Db 652 RLEKTF-EYIDTF--HKFKMGVS--GCPRSVCSG-VKDFGIIISVEN-GF-----QIY 697
Qy 449 I--KGMVVFSEMGDSNASVPTQPV-----YREMFHGHGKAKPDTISITFSKVA 497
Db 698 IGGNGGTEV--EKAEFLTTTETEDDEVIKLCGALMQYRET-GIYAE--TA-PWLRLRG 750
Qy 498 YENGVEKELGL--ERQVLFPVKNCRNITKDP-KFNDKTAKITVDPKTFEYFVDGKLCISK 554
Db 751 FEN-VKEVL-LDPERQ-----N-EL-----FERIMD--AKKAVEAPWEAI-----TS- 788
Qy 555 PTSQVPLAQR 564
Db 789 -NAQ-----ARR 794

RESULT 15

US-10-505-928-827

; Sequence 827, Application US/10505928

; Publication No. US20060088532A1

; GENERAL INFORMATION:

; APPLICANT: Ludwig Institute for Cancer Research et al.

; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES

; FILE REFERENCE: 28967/39178

; CURRENT APPLICATION NUMBER: US/10/505,928

; CURRENT FILING DATE: 2004-08-27

; PRIOR APPLICATION NUMBER: US 60/363,019

; PRIOR FILING DATE: 2002-03-07

; NUMBER OF SEQ ID NOS: 866

; SOFTWARE: PatentIn 3.2

; SEQ ID NO 827

; LENGTH: 4834

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-505-928-827.

Query Match

Best Local Similarity 10.1%; Score 303; DB 6; Length 4834;

Matches 157; Conservative 117; Mismatches 212; Indels 201; Gaps 85;

Qy 8 YNVTYGPTKGDKVRIGDGLDPLA--E-VHEDYTYTG-----LWAEVHEHY-----TTYGEELKFG 45

Db 4007 YATGYG--AGGRIGIGGTESVSTPTLLESIQHFVFIKKVAVNSGGKCHLALSSEGEVYSWG 4064

Qy 46 AGKTTREG-MGQSN-SP-DE-NTLDELVTNMIIDYT--GIYKADI---G-IKN-GKIHG 94

Db 4065 EAE---DGKLGHNRSPCDRPRVIE-SLRGIEVVDVAAGGAHSACTAAGDLTYWGK--G 4118

Qy 95 -IGKAGNKDMQDGVSPHVMVVGTEALAGEGM1-ITAGGIDSHTHFLSPQGFPTALANGV 152

Db 4119 RVGRLGHSDESQOLKPKLV-----EALQHRVVDIACGSDAQTLCLTDDDTVMSWGDGD 4173

Qy 153 TTMFG-GGT-G---P--VD---GTNATTITPG-KWNLHRLRAAEYSNMVNG-F--LGKG 198

Db 4174 YGKLGSGSDCKVPMKIDSITGLGVKVEGSGQFSV-ALTKSGAVYTWGKGDYHRLGHG 4232
Qy 199 NSSKKQLVEQVEAGATGFKLHEDWGTTTPSAIDHCLSVADYDQVCIHT--DTVNEAGY 256
Db 4233 SDDHVR--PROVQ-GLQCKKVIAT--ATGSL--HCVCCTE--DGEV--YTWGD--NDEGQ 4280
Qy 257 V-DDTLNAM-NGRAIHAYHIEG-----ACGHSFDPVITMAGELNLPSSST--TPT-IP 304
Db 4281 LGDGTNAIORPRLVAA--LQKKKNRVACGSAH-----TLAWSTS--KPASAGKLPAQVP 4332
Qy 305 YTINTVAE---HL-DMLMTCHLIDKR---I-REDLOFS--QSRIRPGSIAAEDVLHDM 352
Db 4333 MEYNHLQEIPIIALNRNLLHLLHLSLFCPCIPMFDEGLSDETGPG-SVGF-DTLR-- 4388
Qy 353 GVIAWTSDS-----QA-MGR---AGEVIP--RTWQTADKNNKEFGKL--PEDGKDNDN 398
Db 4389 GILISQGEAAFRKVVQATWVRDRQHGFPVELNRI-QV--KRSRSKGLAGP-DG-TKSV 4443
Qy 399 F-RIKRYISKYTNP---ALTHGV--SEYIG-SVEE---G---KIADLV--VWNPAPFGV 443
Db 4444 FGQWCAKMSF--GPDSELLLPHRYWVKVFGESVDDCGGYSSESIAETCBELQN---GL 4497
Qy 444 KPKIVI-KGGWVFESEMGDSNASV---PTPQ-PVYVREMGHGHGAKFDTSTITFVSKVA 497
Db 4498 TPLLIIVTPNR---DESG-ANRDCYLLSPAARAPV-HSSMF-----RF---LGVLLGIA 4543
Qy 498 YENGVEKELGL-E---RQV---LPVKNCRNITKKDF-----KFNDKTAKITVDPKTF 542
Db 4544 IRTGSPLSLNLAEFPVKQLAGMSLTADLSEV-DKDFIPGLMYIRDNEATS-----BEF 4596
Qy 543 EVFVDGKCTSKPTS--Q-VPLAQRYT 566
Db 4597 EAM---SLPPTVPSASGQDIQLSSKHT 4620

Search completed: August 9, 2006, 21:48:34
Job time : 36 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:38:07 ; Search time 41 Seconds
(without alignment)
1332.954 Million cell updates/sec

Title: US-09-904-994B-3

Perfect score: 2999

Sequence: 1 MKMKQRYVNTYGTGKGVK.....KLCTSKPTSQVPLAQRVTF 568

Scoring table:

BLOSUM62
Gapop 4.0 , Gapext 1.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.80.*
2: PIR1.*
3: PIR2.*
4: PIR3.*
5: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2297	76.6	569	1 URKCBP	urease (EC 3.5.1.5)
2	2244	74.8	569	1 S35291	urease (EC 3.5.1.5)
3	2044	68.2	569	2 C36950	urease (EC 3.5.1.5)
4	2027	67.6	571	2 F83681	urease (alpha subu
5	1998	66.6	569	2 D69729	urease (alpha subu
6	1942	64.8	569	2 S75169	urease (EC 3.5.1.5)
7	1916	63.9	568	2 AG2264	urease (alpha chain
8	1900	63.4	569	2 S47104	urease (EC 3.5.1.5)
9	1849	61.7	568	2 C75586	urease, alpha subu
10	1843	61.5	572	2 H64075	urease (EC 3.5.1.5)
11	1838	61.3	566	2 H83037	urease (alpha subu
12	1838	61.3	570	2 S42607	urease (EC 3.5.1.5)
13	1834	61.2	567	1 C36138	urease (EC 3.5.1.5)
14	1826	60.9	567	1 D43719	urease (EC 3.5.1.5)
15	1822	60.8	567	1 S08480	urease (EC 3.5.1.5)
16	1818	60.6	568	2 G85654	probable urease st
17	1818	60.6	568	2 D90794	urease (alpha subu
18	1814	60.5	569	2 A97648	urease (alpha chain
19	1814	60.5	569	2 AG2871	urease (alpha subu
20	1805	60.2	838	2 A96699	probable urease F1
21	1801	60.1	570	1 Ar3458	urease (EC 3.5.1.5)
22	1797	59.9	571	2 E90027	urease (alpha subu
23	1795	59.9	572	2 AC0325	urease (EC 3.5.1.5)
24	1792	59.8	571	2 S38485	urease (EC 3.5.1.5)
25	1770	59.0	840	1 URJB	urease (EC 3.5.1.5)
26	1734	57.8	573	2 A13332	urease (EC 3.5.1.5)
27	1733	57.8	835	2 T37939	urease (EC 3.5.1.5)
28	1721	57.4	598	2 G82890	urease complex com
29	1717	57.3	573	2 S36028	urease (EC 3.5.1.5)

30	1624	54.2	577	2 B70665	probable ureC prot
31	1549	51.7	614	1 S10032	urease (EC 3.5.1.5)
32	1266	42.2	558	2 T29056	urease (EC 3.5.1.5)
33	1222	40.7	569	2 T50711	urease (EC 3.5.1.5)
34	441	14.7	129	2 T08993	urease (EC 3.5.1.5)
35	339	11.3	961	2 AD0548	pyruvate autotransp
36	333	11.1	1078	2 F97907	hyaluronate lyase
37	332	11.1	424	2 H96963	dihydroorotase [im
38	331	11.0	1066	2 B95037	hyaluronidase [imp
39	329	11.0	1983	2 G86643	hypothetical prote
40	328	10.9	579	2 AD1664	adenine deaminases
41	326	10.9	897	2 AD2906	aconitate hydratase
42	326	10.9	897	2 E97681	aconitate hydratase
43	325	10.8	526	2 B75202	d-aminoacylase (as
44	324	10.8	2020	2 C48399	ABC-type transport
45	324	10.8	2198	2 T20371	hypothetical prote

ALIGNMENTS

RESULT 1

URKCBP

urease (EC 3.5.1.5) 62K chain - Helicobacter pylori (strains 26695, J99, and others)
N:Alternate names: urease beta chain; urease chain B; urease large subunit
C:Species: Helicobacter pylori
C:Date: 30-Sep-1991 #sequence revision 02-Dec-1994 #text change 05-Oct-2004
C:Accession: B38537; H64528; A71977; F41834; S07885; S12487; A49215; B61371; A35306; B41:
R:Labigne, A.; Cussac, V.; Courcoux, P.
J. Bacteriol. 173, 1920-1931, 1991

A:Title: Shuttle cloning and nucleotide sequences of Helicobacter pylori genes responsibl
A:Reference number: A38537; MUID:91161505; PMID:2001995
A:Accession: B38537
A:Molecule type: DNA

A:Residues: 1-569 <LAB>

A:Cross-references: UNIPROT:P14917; UNIPARC:UPI000002C20A; GB:M60398; NID:g149007; PIDN:?
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64528

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-569 <TOM>

A:Cross-references: UNIPARC:UPI000002C20A; GB:AE000529; GB:AE000511; NID:g2313152; PIDN:?
A:Experimental source: strain 26695
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J
Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: A71977

A:Molecule type: DNA

A:Residues: 1-569 <ARN>

A:Cross-references: UNIPARC:UPI000002C20A; GB:AE001446; GB:AE001439; NID:g4154573; PIDN:?
A:Experimental source: strain J99
R:Cussac, V.; Ferrero, R.L.; Labigne, A.
J. Bacteriol. 174, 2466-2473, 1992

A:Title: Expression of Helicobacter pylori urease genes in Escherichia coli grown under i
A:Reference number: A41834; MUID:92210488; PMID:1313413
A:Accession: F41834

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 565-569 <CUS>

A:Cross-references: UNIPARC:UPI000002D2AD; GB:M84338
R:Clayton, C.L.; Pallen, M.J.; Kleanthous, H.; Wren, B.W.; Tabaqchali, S.
Nucleic Acids Res. 18, 362, 1990

A:Title: Nucleotide sequence of two genes from Helicobacter pylori encoding for urease st
A:Reference number: S07884; MUID:90221820; PMID:2326167

A:Accession: S07885

A:Molecule type: DNA
A:Residues: 1-9,'A',11-18,'A',20-103,'T',105-180,'F',182-192,'F',194-217,'L',219-272,'Y'
A:Cross-references: UNIPARC:UPI0000172DA6; EMBL:X17079
R:Clayton, C.L.
Submitted to the EMBL Data Library, October 1989
A:Reference number: S12487
A:Accession: S12487
A:Molecule type: DNA
A:Residues: 1-9,'A',11-103,'T',105-180,'F',182-192,'F',194-217,'L',219-272,'Y',274-539,'
A:Cross-references: UNIPARC:UPI000002DB18; EMBL:X17079; NID:g43633; PIDN:CAA34933.1; PID
R:Turbet, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.
Infect. Immun. 60, 5259-5266, 1992
A:Title: Purification and characterization of the urease enzymes of Helicobacter species
A:Reference number: A49215; MUID:93084378; PMID:1452359
A:Accession: A49215
A>Status: preliminary
A:Molecule type: Protein
A:Residues: 1-12 <TUP>
A:Cross-references: UNIPARC:UPI00000B8645; PIDN:AAB24355.1; PID:g260964
A:Experimental source: strain NCTC 11637
R:Evans Jr., D.J.; Evans, D.G.; Kirkpatrick, S.S.; Graham, D.Y.
Microb. Pathog. 10, 15-26, 1991
A:Title: Characterization of the Helicobacter pylori urease and purification of its subu
A:Reference number: A61371; MUID:93112104; PMID:1857197
A:Accession: B61371
A:Molecule type: Protein
A:Residues: 1-21,'C',23-26,'C',28,'I',30 <EVA>
A:Cross-references: UNIPARC:UPI0000172DA7
R:Dunn, B.E.; Campbell, G.P.; Perez-Perez, G.I.; Blaser, M.J.
J. Biol. Chem. 265, 9464-9469, 1990
A:Title: Purification and characterization of urease from Helicobacter pylori.
A:Reference number: A35306; MUID:90264448; PMID:2188975
A:Accession: A35306
A:Molecule type: Protein
A:Residues: 1-20 <DUN>
A:Cross-references: UNIPARC:UPI00000B3DE0
R:Hu, L.T.; Mobley, H.L.T.
Infect. Immun. 58, 992-998, 1990
A:Title: Purification and N-terminal analysis of urease from Helicobacter pylori.
A:Reference number: A41502; MUID:90202165; PMID:2318539
A:Accession: B41502
A:Molecule type: protein
A:Residues: 1-5,'A',7-15 <HUA>
A:Cross-references: UNIPARC:UPI0000172DA8
C:Genetics:
A:Gene: ureB; HP0072
C:Complex: heterodimer of 26K chain (see PIR:URKCAP) and 62K chain
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and tw
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterodimer; hydrolase; metalloprotein; nickel
F:4-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
F:219,248,274/Binding site: nickel 1 (Lys, His, His) #status predicted
F:219/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predict
F:221,322/Active site: His #status predicted
Query Match 76.6%; Score 2297; DB 1; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.9e-46;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKOEYVNTYPTKDGKVRIGDGLDIAEVEHDYTTTGELKFGAGKTIREGMGQSNPSD 61
DB 3 KISRKEYVSMYPTTKDGKVRIGDGLDIAEVEHDYTTTGELKFGAGKTIREGMGQSNPS 62
QY 62 ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMQDGVSPHVVGVGTEALA 121
DB 63 KEELDLIITNALIVDYTYGIYKADIGIKGKIAGIKGNGKQMDGVKNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOFPPTALANGVTTMTFGGGTGPVDGNTATTITPGKNLHRM 181
DB 123 GEGLIIVTAGGIDTHIHFIPOQIPTAFASGVTTMTGGGTGPADGTNATTITPGRNLRKM 182

QY 182 LRAAEYSNNVGFGLKGNSSSKQLVQVEAGAIGPKLHEDWGTTPSAIDHCLSVADEYD 241
DB 183 LRAAEYSNNLGLFAGKSNASNDASLADQTEAGAIGFKIHEDWGTTPSAINHALVDADKYD 242
QY 242 VQVCIHDTTVNEAGYVDDTLNAMNGRAHYHIEGAGGGHSPDVTIMAGELNLPSTTP 301
DB 243 VQVAIHTDNLNAGCVEDTMAAIGRTMTFTTEGAGGGHAPDIIIKVAGHNLIPASTNP 302
QY 302 TTPYTTINTVAEHLMLMTCHHLDRKIREDLQFSQSRIRPGSIAAEVDLHDMGVIAMTSSD 361
DB 303 TTPFTVNTAEHMDMLWCHHLDKSKEVDQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SQAMGRAGVIRPTWTQADNKKKFGKLPEDGKDNDFRIKRYISKYTINPALTHGVSEY 421
DB 363 SQAMGRVGEVIRPTWTQADNKKKFGKLPEDGKDNDFRIKRYISKYTINPALTHGVSEY 422
QY 422 IGSVEGKIADLVNVPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYVYREMGHH 481
DB 423 VGSVEVGKADLVNVPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYVYREMGFAHH 482
QY 482 GKAKPDTSTTFVSKVAYENGVEKELGLERQVLFPVKNCRNITKKDFKFNDKTAKITVDPKT 541
DB 483 GKAKYDANITFVSOAYDKGIEKELGLERQVLFPVKNCRNITKKDMQFNDDTAHIEVNPET 542
QY 542 FEVFDGKLCSTKPTSQVPLAQRYPFF 568
DB 543 YHVFVDGKGVTSKPAKVSIAQLFSIF 569
RESULT 2
S35291
A:Alternate names: urease beta chain
N:Alternate names: Helicobacter felis
C:Species: Helicobacter felis
C:Date: 31-Dec-1993 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004
C:Accession: S35291
R:Ferrero, R.L.; Labigne, A.
Mol. Microbiol. 9, 323-333, 1993
A:Title: Cloning, expression and sequencing of Helicobacter felis urease genes.
A:Reference number: S35290; MUID:94018627; PMID:8412683
A:Accession: S35291
A:Molecule type: DNA
A:Residues: 1-569 <PER>
A:Cross-references: UNIPROT:Q08716; UNIPARC:UPI000002F34B; EMBL:X69080; NID:g396160; PIDN
C:Genetics:
A:Gene: ureB
C:Complex: heterodimer
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and tw
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterodimer; hydrolase; metalloprotein; nickel
F:1-569/Product: urease 62K chain homology <U62>
F:4-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
F:219,248,274/Binding site: nickel 1 (Lys, His, His) #status predicted
F:219/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predict
F:221,322/Active site: His #status predicted
Query Match 74.8%; Score 2244; DB 1; Length 569;
Best Local Similarity 72.5%; Pred. No. 3.5e-45;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;
QY 2 KMKOEYVNTYPTKDGKVRIGDGLDIAEVEHDYTTTGELKFGAGKTIREGMGQSNPSD 61
DB 3 KISRKEYVSMYPTTKDGKVRIGDGLDIAEVEHDYTTTGELKFGAGKTIREGMGQSNPS 62
QY 62 ENTLDLVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMQDGVSPHVVGVGTEALA 121
DB 63 SYELDLVLTNALIVDYTYGIYKADIGIKGKIAGIKGNGKQMDGVNDNLCVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOFPPTALANGVTTMTFGGGTGPVDGNTATTITPGKNLHRM 181

Db 123 AEGLVITAGGIDTHIFISPOQIPTAFASGVTTMIGCGTPADGTNATTITPGRANLXSM 182
Qy 182 LRAAEYSVMVGFYKGNSSKKQLVEAGATGFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAAEYAMNGLFAGKGVSEPSLSDQIEAGATGFKLHEDWGTTPAAIHCLNVADEYD 242
Qy 242 VQVCITHTVNEAGVDDTLNAMNGRAIYAHIEGAGGCHGSPDVTWAGELNLPSSSTTP 301
Db 243 VQVAIHDTLNEAGCVDTLEAIGRIHTHTHTEGAGGCHGAPDVKMAGEFNILPASTNP 302
Qy 302 TTPYTINTVAEHLDMLTCHHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVAMTSSD 361
Db 303 TTPKTNTAEAEHMDLVMVCHLDSIKEDVQFADSRIRPQTIAEDQLHDMGIFSISSD 362
Qy 362 SQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRPKRYISKYITINPALTHGVSEY 421
Db 363 SQAMGRVGEVIRTWQTADKNKKEFGKLPEDGKNDNFRPKRYISKYITINPQIAHGISDY 422
Qy 422 IGSVEEGKIADLVVNNPAFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFGHH 481
Db 423 VGSVEVGKIADLVVNSPAPFGKINMIIKGGFIALSQMGDANASITPTPQPVYVREMFGHH 482
Qy 482 GKAKFDTISITFVSKVAYGVGKELGLERQVLPVKNCRNITTKDPKFNDKTAKITVDPKT 541
Db 483 GKAKFDTNITFVSOAYKAGIKKEGLDRAAPPVKNCRNITTKDLKFNVDVTAHIDVNPET 542
Qy 542 FEVVDGKLCSTKPTSOVPLAQRVTFP 568
Db 543 YKVKVDGKEVTSKAADELSLAQLNLF 569

RESULT 3
C36950
Urease (EC 3.5.1.5) 61k chain - Bacillus sp. (strain TB-90)
N;Alternate names: ureC protein
C:Species: Bacillus sp.
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 05-Oct-2004
R;Maeda, M.; Hidaka, M.; Nakamura, A.; Masaki, H.; Uozumi, T.
J. Bacteriol. 176, 432-442, 1994
A;Title: Cloning, sequencing, and expression of the thermophilic Bacillus sp. strain TB-90 u
A;Reference number: A36950; MUID:94117379; PMID:8288539
A;Accession: C36950
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-569 <MAE>
A;Cross-references: UNIPROT:Q07397; UNIPARC:UPI0000137D67; GB:D14439; NID:G393296; PIDN:
C;Superfamily: urease, alpha subunit; urease 62K chain homology
F;5-552/Domain: urease 62K chain homology <U62>

Query Match 68.2%; Score 2044; DB 2; Length 569;
Best Local Similarity 66.4%; Pred. No. 2.2e-40;
Matches 378; Conservative 76; Mismatches 108; Indels 7; Gaps 4;

Qy 3 MKKQEVNTYPTKDGKVRGLDITLWAEVEHDYTYGEEELKFGAGKTIREGMQO---SNS 59
Db 5 MSRKQYADMFGPTVGDALRLADSELFTIEKDYTYGDEVKFGGKVRDGMGQHPPLATS 64
Qy 60 PDENTLDTVTNAMIIDYTYGKADIGIKNGKIHGIGKAGNKMDDGVSPHVMVGVGTEA 119
Db 65 -DE-CVDLVLVTNAIIVDTYGYKADIGIKDGMIAISIGKAGNPLMDGVD--MVIGAAATEV 120
Qy 120 LAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGTPGVDTGNATTITPGKNLH 179
Db 121 IAAEGMIITAGGIDAHITFCPOQIEFALASGVTTMIGGGTGPATGTNATTCTPGPMNIH 180
Qy 180 RMLRAAEYSNMVGFYKGNSSKKQLVEAGATGFKLHEDWGTTPSAIDHCLSVAD 239
Db 181 RMLQAAEFINLGFYKGNCSDEAPLKEQIEAGAVGLHEDWGSTAAADITCLKVADR 240
Qy 240 YDVQVCIHTDVNEAGVDDTLNAMNGRAIYAHIEGAGGCHGSPDVTWAGELNLPSSST 299

Db 241 YDVQVAIHDTLNEGGFVBDTLKAIDGRVHTYHTEGAGGCHAPDIIKAAAGFPNLPSSST 300
Qy 300 TTPYTINTVAEHLDMLTCHHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVAMTSS 359
Db 301 NTRPTTYINTLEHLDMLVMVCHLDAINIPEDIAFADSRIRKETIAAEDVLHDLGVFSMIS 360
Qy 360 SDSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRPKRYISKYITINPALTHGV 419
Db 361 SDSQAMGRVGEVIRTWQTADKNKKEFGKLPEDGKNDNFRPKRYISKYITINPAIAHGA 420
Qy 420 EYIGSVEEGKIADLVVNNPAFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFG 479
Db 421 DVGVSVEVGKIADLVVNNPAFGVKDELVLKGGMIAVSTWGDNPASITPTPQVLYRPMFA 480
Qy 480 HGGKAKFDTISITFVSKVAYGVGKELGLERQVLPVKNCRNITTKDPKFNDKTAKITVDP 539
Db 481 AKGDGATQTSITFVSKAAVEKGIHEQLGKVKVPHGIRKLTKKOILLNDKTPKIDVDP 540
Qy 540 KTFEVPVVDGKLCSTKPTSOVPLAQRVTFP 568
Db 541 QTYEVKVDGQLVTCBPAPETVPMQAORYPLF 569

RESULT 4
P83681
Urease alpha subunit ureC [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 05-Oct-2004
C:Accession: F83681
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F83681
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-571 <STO>
A;Cross-references: UNIPARC:UPI00000C3801; GB:AP001507; GB:BA000004; NID:gl0172612; PIDN:
A;Experimental source: strain C-125
C;Genetics:
A;Gene: ureC
C;Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 67.6%; Score 2027; DB 2; Length 571;
Best Local Similarity 64.4%; Pred. No. 5.8e-40;
Matches 368; Conservative 88; Mismatches 112; Indels 3; Gaps 2;

Qy 1 MKKQEVNTYPTKDGKVRGLDITLWAEVEHDYTYGEEELKFGAGKTIREGMQO--N 58
Db 1 MKLTRAQHSLYGPVGVKVRGLADTLLLEIEKDYTYGDEVKFGGKVRDGMGQSAVY 60
Qy 59 SPDENTLDTVTNAMIIDYTYGKADIGIKNGKIHGIGKAGNKMDDGVSPHVMVGVGTE 118
Db 61 TRDEGVLDLIITNATIIDYTYGKADIGIKNGHIVGIGKGNPDIMDGVESHMIVGASTE 120
Qy 119 ALAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFGGTPGVDTGNATTITPGKNL 178
Db 121 AIAGEGLVITAGGIDAHITFISPOQIDVAIASSITMTLGGTGPATGTATTCTPGKWI 180
Qy 179 HRMLRAAEYSNMVGFYKGNSSKKQLVEAGATGFKLHEDWGTTPSAIDHCLSVAD 238
Db 181 ERMLEAADAFPNLGFYKGNASTPAPLREQIEAGATGFKLHEDWGTTPAAITCLSVAD 240
Qy 239 EYDVQVCIHTDVNEAGVDDTLNAMNGRAIYAHIEGAGGCHGSPDVTWAGELNLPSS 298
Db 241 RMDVQVAIHTDTLNEAGFVBDTIIKAGDRVHTYHTEGAGGCHAPDIIKAAAGFPNLPSS 300
Qy 299 TTPYTINTVAEHLDMLTCHHLDRKIREDLQFSOSRIRPGSIAAEDVLHDMGVAMT 358
Db 301 TNPTPTVTNIDEHLDMLVMVCHLDPNVDPEDVAFADSRIRPETIAAEDILQDLGVSMI 360
Qy 359 SSDSQAMGRAGEVIRPTWQTADKNKKEFGKLPED--GKNDNFRPKRYISKYITINPALTHG 417

Db 361 SSDSQMGRVGEVIIRTWOTADKMKKORGALAEADQKGNVDNRIRKRYVSKYTINPAITHG 420
Qy 418 VSEYIGSVBEGKIADLVVWNPFAFGVGPVKIVIKGMVVFSEMGSDNASVPTPQPVPYREM 477
Db 421 IDYVGSVEVGLADLVLDWPRPFGVPELILKGLIAYSQMGDPNASIPTPQVPFSRPM 480
Qy 478 FGHGKAKFDTSTIFVSKVAYENGVEKGLGLERQVLVVKNCNITKKDFKNDTKAKITV 537
Db 481 FGSFGRAGRTTSTIFLSKAAMDGLVGHGALGLQKIAHVKNCRSIKRSKMYNDATPNIEI 540
Qy 538 DPKTFFVFDGKLCSTKPTSOVPLAQRYPF 568
Db 541 DPEYEVKVDGEMITCAPFBEVALAQRYPF 571

RESULT 5
D69729
urease (alpha subunit) ureC - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Oct-2004
C:Accession: D69729
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D69729
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-569 <KUN>
A:Cross-references: UNIPROT:P77837; UNIPARC:UPI0000060894; GB:Z99122; GB:AL009126; NID:9
A:Experimental source: strain 168
C:Genetics:
A:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
F:3-552/Domain: urease 62K chain homology <U62>

Query Match 66.6%; Score 1998; DB 2; Length 569;
Best Local Similarity 62.7%; Pred. No. 2.9e-39;
Matches 359; Conservative 94; Mismatches 111; Indels 9; Gaps 6;
Qy 1 MKMKQBYVNTYGTGDKVRLGDTDLMAVEVHDYTYGEEKFGAGKTIREGMSQSN-- 58
Db 1 MKMSREYAEFLGPTGDKIRLGDITLWIEVEKDPVYGEEMI FGGGKTIRDMGQNGRI 60
Qy 59 SPDENTLDVITNAMIIDYTGIVKADIGIKNGKTHGKAGNKMDCQGVSHMVGVGTE 118
Db 61 TKGKDALDVLITNVLLDYTGIVKADVGKDRIVGVGKGNPDIMDGVDPHVMVIGAGTE 120
Qy 119 ALAGEGMIITAGGIDSTHFLSPQFPALANGVTTMFGGGTGPVDTGNATTITPGKNL 178
Db 121 VISSEGLITAGGVDTTHIFICPQOMEVALSSGVTTLLGGTGATGSKATCTTSGAWM 180
Qy 179 HRMLRAAEYSVMVGFGLKGNSSKKQIVQEVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 238
Db 181 ARMLEAAEEFPINVGFLKGNASDKAPLIEQVEAGAIGLKLHEDWGTTPSAIKTCMEVVD 240
Qy 239 EYDVQVCIHTDVTNEAGVDDTLNMGRAIHAHYIEGAGGCHSPDVTMAGELNIPSS 298
Db 241 EADIQVAIHTDVTNEAGVDDTLNMGRAIHAHYIEGAGGCHSPDVTMAGELNIPSS 300
Qy 299 TTPTIPTINTVAEHLDMLTCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAVT 358

Db 301 TTPTIPTINTVAEHLDMMTCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAVT 360
Qy 359 SSDSQMGRVGEVIIRTWOTADKMKKORGALAEADQKGNVDNRIRKRYVSKYTINPAITHG 417
Db 361 SSDSQMGRVGEVIIRTWOTADKMKKORGALAEADQKGNVDNRIRKRYVSKYTINPAITHG 418
Qy 418 VSEYIGSVBEGKIADLVVWNPFAFGVGPVKIVIKGMVVFSEMGSDNASVPTPQPVPYREM 477
Db 419 LSHEVGSVEKGLADLVLDWPRPFGVPELILKGLIAYSQMGDPNASIPTPQVPFSRPM 478
Qy 478 FGHGKAKFDTSTIFVSKVAYENGVEKGLGLERQVLVVKNCNITKKDFKNDTKAKITV 537
Db 479 YASYGKANRSTSTIFMSQASIERGVAESLGEKRIKSPVKIRKLSKLDKMLNSALPKIEI 538
Qy 538 DPKTFFVFDG-KL-CTSQTPOVPLAQRYPF 568
Db 539 DPKTFFVFDGEEELSC--QFVDYVPLGQRYFLF 569

RESULT 6
S75169
urease (EC 3.5.1.5) alpha chain - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sl11750
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Accession: S75169
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75169
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-569 <KAN>
A:Cross-references: UNIPROT:P73061; UNIPARC:UPI0000137D79; EMBL:D90903; GB:AB001339; NID:
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase; metalloprotein; nickel
F:5-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted

Query Match 64.8%; Score 1942; DB 2; Length 569;
Best Local Similarity 62.6%; Pred. No. 6.3e-38;
Matches 356; Conservative 96; Mismatches 112; Indels 5; Gaps 4;
Qy 2 KMKQBYVNTYGTGDKVRLGDTDLMAVEVHDYTYGEEKFGAGKTIREGMSQSN-SP 60
Db 4 RMDREYAHFTGPTGDKVRLADTLFTEVEQDYATYGVDEKFGGKVRDGMGQSPLSR 63
Qy 61 DENTLDVITNAMIIDYTGIVKADIGIKNGKTHGKAGNKMDCQGVSHMVGVGTEAL 120
Db 64 AEGAVDVVITNALIDWGVIVKADVGIKNGRIYAGKAGNPHIQDNVS--IIIGPSTEAI 121
Qy 121 AGEEMIITAGGIDSTHFLSPQFPALANGVTTMFGGGTGPVDTGNATTITPGKNLHR 180
Db 122 AGEEMILTAGGIDAHVHFICPQIETALASGVTTLVGGGTGPAAGTKATCTTCGANNIHR 181
Qy 181 MURAAEYSVMVGFGLKGNSSKKQIVQEVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240
Db 182 MLQAADGFPINLGFGLKGNQSPAALAEQIKAGAILKLHEDWGTTPPAIDNCLGVAEDY 241
Qy 241 DVQVCIHTDVTNEAGVDDTLNMGRAIHAHYIEGAGGCHSPDVTMAGELNIPSS 300
Db 242 DVQVCIHTDVTNEAGVDDTLNMGRAIHAHYIEGAGGCHSPDVTMAGELNIPSS 301
Qy 301 PTPTIPTINTVAEHLDMLTCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAVT 360

Db 302 PTRPYTINTLBEHLDMVMCHLHNRNIPEDVAFABESRIRRETIAAEDILHDLGAFSISS 361
Qy 361 DSQAMGRAGEVLPRTWQTADKNKKEFGKLP-EDGKNDNFRKRYISKYTTINPALTGHVS 419
Db 362 DSQAMGRAGEVLPRTWQTAHKMKVQKQGLPOSTG-NNDNLRAKRYVAKYTTINPAITHGIS 420
Qy 420 EYIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 479
Db 421 BEIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 480
Qy 480 HHGKAKFDTSTITFVSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDKTAKITVD 539
Db 481 NYGGAIAATSVTFVSKAAKADIGKGLSKPLVAVKNIRQLTKRDLKLNLYLPHIEVD 540
Qy 540 KTFEYFVDGKLTCKTSPKTSQVPLAQRITFF 568
Db 541 EYEVADGELLTCEPASVLPMAQRYFLF 569
RESULT 7
AG2264
urease alpha chain [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 05-Oct-2004
C:Accession: AG2264
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2264
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-568 <KUR>
A:Cross-references: UNIPARC:UPI00000CE887; GB:BA000019; PIDN:BA075369.1; PID:gl7132803;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3670
C:Superfamily: urease, alpha subunit; urease 62K chain homology
Query Match 63.9%; Score 1916; DB 2; Length 568;
Best Local Similarity 61.4%; Pred. No. 2.6e-37;
Matches 350; Conservative 95; Mismatches 117; Indels 8; Gaps 4;
Qy 2 KMKKQEVNTVPTKGDVRLGDTDLWAEVEHDYTTVGEELKFGAGKTIREGMGQS---N 58
Db 4 RMRQVAYETVPTVGRIRLADTELFQVEQDFTTYGDEVNFGGKVRDGMGQSPIAN 63
Qy 59 SPDENTLVLVTNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDQVSPHVVVGTE 118
Db 64 A--DGAVDLVTNALILDWGVKADIGIKDGKIFKIGKAGNPYIQHDV--IIIGPTE 119
Qy 119 ALAGEGMIITAGDISHTHFLSPQFPPTALANGVTVMFGGTPGVDGNTATTTPGKNL 178
Db 120 ALAGEGMIITAGDISHTHFLSPQFPPTALANGVTVMFGGTPGVDGNTATTTPGKNL 179
Qy 179 HRMLRAAEYSNMVGFCKGNSKKQVQVEAGATGFKLHEDWGTTPSAIDHCLSVAD 238
Db 180 YRMLQADAPFPMNLGFLGKNSAQVQGLVQIFAGATGLKHLHEDWGTTPATIDTCLTVAD 239
Qy 239 EYDVQVCHTDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNLPSS 298
Db 240 EYDVQVAIHDTLNEAGFVEDTAAFKNRAIHYHIEGAGGSHADPDIKVCQANVLPSS 299
Qy 299 TTPTIPYNTVAEHLDMTCHHLDKRIREDLQFSQSRIRPGSIRPGSIAAEDVLHDMGVAMT 358
Db 300 TNTPRPYNTVAEHLDMTCHHLDKRIREDLQFSQSRIRPGSIRPGSIAAEDVLHDMGVAMT 359
Qy 359 SSDSQAMGRAGEVLPRTWQTADKNKKEFGKLPEDGKNDNFRKRYISKYTTINPALTGHV 418
Db 360 ASDSQAMGRAGEVLPRTWQTSHKMKVQKQGLTGA--EADNLRAKRYVAKYTTINPAITHGI 419

Qy 419 SEYIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 478
Db 419 AQVGVSEAGKADLCWNPAPFGVKPEIVIKGMTAWSONGDANASIPITPOVHMPMP 478
Qy 479 GHGKAKFDTSTITFVSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDKTAKITVD 538
Db 479 GSFAGARNATSLTFVSSQAALEREIPQQLGLRKSAAVAVSGTRQLTKQDMKLDALPHIEVD 538
Qy 539 KTFEYFVDGKLTCKTSPKTSQVPLAQRITFF 568
Db 539 SESYEVRADGELLTCEPASVLPMAQRYFLF 568
RESULT 8
S47104
urease (EC 3.5.1.5) 62K chain - Bacillus pasteurii
N:Alternate names: ureC protein
C:Species: Bacillus pasteurii
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S47104
R:Moersdorf, G.; Weimann, P.; Kaltwasser, H.
submitted to the EMBL Data Library, June 1994
A:Description: Nucleotide sequence of three genes on a urease encoding DNA-fragment from
A:Reference number: S47102
A:Accession: S47104
A:Molecule type: DNA
A:Residues: 1-569 <MOE>
A:Cross-references: UNIPROT:P41020; UNIPARC:UPI0000137D66; EMBL:X78411; NID:g498708; PID:
C:Genetics:
A:Gene: ureC
A:Start codon: GTG
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterodecamer; hydrolase; metalloprotein; nickel
F:3-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, Lys, Asp) #status predicted
Query Match 63.4%; Score 1900; DB 2; Length 569;
Best Local Similarity 59.1%; Pred. No. 6.4e-37;
Matches 337; Conservative 112; Mismatches 118; Indels 3; Gaps 2;
Qy 1 KMKKQEVNTVPTKGDVRLGDTDLWAEVEHDYTTVGEELKFGAGKTIREGMGQS-- 58
Db 1 MKINRQVAYETVPTVGRVRLADTDL-GEVERDYTLGDEVNFGGKVRDGMGNGTY 59
Qy 59 SPDENTLVLVTNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDDQVSPHVVVGTE 118
Db 60 TRTENVDLLLTNALILDYTIYKADIGYKQGVYIGKGNPDINDGVTPNNIVGTATE 119
Qy 119 ALAGEGMIITAGDISHTHFLSPQFPPTALANGVTVMFGGTPGVDGNTATTTPGKNL 178
Db 120 VIAAEGKIVTAGDITHVHFINPDQVDVALANGITTLFGGCTGPARGSKATVTPGPWNI 179
Qy 179 HRMLRAAEYSNMVGFCKGNSKKQVQVEAGATGFKLHEDWGTTPSAIDHCLSVAD 238
Db 180 EKMLKSTEGLPINVGILGKHGSSAPINEQDAGAAGLKIHEDMGATPASIDRSITVAD 239
Qy 239 EYDVQVCHTDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNLPSS 298
Db 240 EADVQVAIHDTLNEAGFLEDITVRAINRVIHSFHVGEAGGSHADPDIIMAGHPNVLPSS 299
Qy 299 TTPTIPYNTVAEHLDMTCHHLDKRIREDLQFSQSRIRPGSIRPGSIAAEDVLHDMGVAMT 358
Db 300 TNTPRPYNTVAEHLDMTCHHLDKRIREDLQFSQSRIRPGSIRPGSIAAEDVLHDMGVAMT 359
Qy 359 SSDSQAMGRAGEVLPRTWQTADKNKKEFGKLPEDGKNDNFRKRYISKYTTINPALTGHV 418
Db 360 STDALAMGRAGEVLPRTWQTADKNKKEFGKLPEDGKNDNFRKRYISKYTTINPALTGHV 419
Qy 419 SEYIGSVBEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYVREMF 478
Db 420 AHEVGSIEBGFADLVLPWEPKPFVGVKADRVIKGGIIAAYIGDPSASITPQVPMGRRMY 479
Qy 479 GHGKAKFDTSTITFVSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKFNDKTAKITVD 538

Db 480 GTVGDLDHNTITFMSKSSIQGGVPKGLKRRIGTVKNCRNIGKMKWMDVTTDIDIN 539
QY 539 PKTEFEVVDGKLGKTSKPTSOVPLAQRVTF 568
Db 540 PETYEVDGEVLTCBPVKELPMAQRVFLP 569

RESULT 9
C75586
u:ase, alpha subunit - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-Oct-2004
C:Accession: C75586
R:White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75586
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-568 <WHI>
A:Cross-references: UNIPARC:UPI00000C1687; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0318
A:Map.position: 2
C:Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 61.7%; Score 1849; DB 2; Length 568;
Best Local Similarity 59.2%; Pred. No. 1.1e-35;
Matches 339; Conservative 100; Mismatches 124; Indels 10; Gaps 7;

QY 1 MMKKQEVNTYPTGKGVRLGDTDLMAEVEHDYTYTGYEELKFGAGKTIREGMQSN- 59
Db 1 MKVSQQVADLYPTGVDRLGDTTELLIEVERDLTYGEEVKFGGKVRIDGLQSSAA 60

QY 60 --PENTLDLVITNAMIIDYGIYKADIGIKNGKIHGKAGNKMDDQGVSPHVVGVGT 117
Db 61 TRDANVPDLVITNALLIDYGVITKADYGVNGKRIISAIGKAGNPGTQGVTPGLTAAST 120

QY 118 EALAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFVGCGTGPVDTGNATTITPGKWN 177
Db 121 EIVAGEGLVLTAGVDVTHIHFIAPQOCWTALESVGTMTMGSGTGPTAGTSATCTPGQWH 180

QY 178 LHRMLRAAEYSVMNVGFLGKGNSSKKQLVQVQVAGAIQFKLHEDWGTPPSAIDH-CLSV 236
Db 181 IHRMLESAGLPLNFGLLGKGNASTQPPPLAQIRAGALGLKLHEDWGTPPAI-HAALSV 239

QY 237 ADEYDVQVCIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPVITMAGELNLP 296
Db 240 AEDYDVQVAIHDTDLNESGFDEAIRAPAGRTIHTFHTEGAGGCHAPDIIRVAGLPNVL 299

QY 297 SSTTPTPTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIA 356
Db 300 SSTNPTPTINTVAEHLDMLVCHHLSPLIPEDVHFAESIRPETIAAEDVLHDMGVFS 359

QY 357 MTSSDSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNPNRIKRYISKYTIINPALTH 416
Db 360 MWSDSQAMGRAGEVIRPTWQAHHKMKVQRGLPADGR-ADNFRARYVAKYTIINPAIAH 418

QY 417 GVSYIGSVGEKGIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQVYRE 476
Db 419 GISHEVSVEVGKGLADLVWSPAFFGAKPSILIKGLVVAQMGDANASITPTPQVYRPP 478

QY 477 MFGHHGKAKFD-TSITFVSKVAYENGVEKGLGLERQVLPVKNCNITKKDFKNDKTAKI 535
Db 479 MFAAYGGCP-DATCLHVFVSQAGLEGHLPDVG--RRYSAVKHTRDIGNKQWLNABTDDI 535

QY 536 TVDPKTEFEVVDGKLGKTSKPTSOVPLAQRVTF 568

RESULT 11

Db 536 QVNPETYEVRVNGBELVTCPEVDDELPLAQKYFLP 568

RESULT 10
H64075
u:ase (BC 3.5.1.5) 62K chain - Haemophilus influenzae (strain Rd KW20)
N:Alternate names: urease alpha chain
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: H64075
R:Plischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.;
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: H64075
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-572 <TIGR>
A:Cross-references: UNIPROT:P44391; UNIPARC:UPI0000137D6B; GB:U32736; GB:L42023; NID:g15:
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase; metalloprotein; nickel
F:5-555/Domain: urease 62K chain homology <U62>
F:141,143,224,367/Binding site: nickel 2 (His, Lys, Asp) #status predicted

Query Match 61.5%; Score 1843; DB 2; Length 572;
Best Local Similarity 59.1%; Pred. No. 1.5e-35;
Matches 338; Conservative 92; Mismatches 136; Indels 6; Gaps 4;

QY 1 MMKKQEVNTYPTGKGVRLGDTDLMAEVEHDYTYTGYEELKFGAGKTIREGMQSNP 60
Db 3 LTISRAQVATYPTGVDKVRIGDTNLWATISQDLLTKGDECKFGGKSVRDMQAQSGTA 62

QY 61 --DE-NTLDLVITNAMIID-YGIYKADIGIKNGKIHGKAGNKMDDQGVSPHVVGVG 116
Db 63 TRDNPVLDVFTTNWMIIDAKLGIKADIGIRGVIGVIGQAGNPDTMDNVTNMMIIGAS 122

QY 117 TEALAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTFVGCGTGPVDTGNATTITPGKW 176
Db 123 TEVNHAGLIATAGGIDTHIHFIQCPQAOHAIESGVTTLLIGSGTGPDGTHATCTPGAW 182

QY 177 NLHRMLRAAEYSVMNVGFLGKGNSSKKQLVQVQVAGAIQFKLHEDWGTPPSAIDHCLSV 236
Db 183 YMERFQAAEALPVNVGFGKNCSTLDPLRQIEAGALGLKLHEDWGATPAVIDSALKV 242

QY 237 ADEYDVQVCIHTDTVNEAGYVDDTLNANGRAIHAYHIEGAGGHSPPVITMAGELNLP 296
Db 243 ADEMDIQAIVHTDLNESGFEDTKAIDGRVHTFHTEGAGGCHAPDIKAAVSNVLP 302

QY 297 SSTTPTPTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIA 356
Db 303 ASTNPTPTPTNTIDHLDMLVCHHLDKRVDPVAFADSRIRPETIAAEDVLHDMGVFS 362

QY 357 MTSSDSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNPNRIKRYISKYTIINPALTH 416
Db 363 INSSDSQAMGRAGEVIRPTWQTADKNKQORGLGNEG--NDNFRIRYAKYTIINPAIAH 420

QY 417 GVSYIGSVGEKGIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQVYRE 476
Db 421 GIAEHIGSLGVKADIADVLWKPMPFGVKPVEVKKGFISYAKMGDPNASIPTPQVYRPP 480

QY 477 MFGHHGKAKFDTSITFVSKVAYENGVEKGLGLERQVLPVKNCNITKKDFKNDKTAKIT 536
Db 481 MYGAQGLATAQAVFVFSQAABKADIRAKFGLHKTETIAVKGCRNVGKGLVHNDVTNIT 540

QY 537 VDPKTEFEVVDGKLGKTSKPTSOVPLAQRVTF 568
Db 541 VDAERYEVRVDGELITCEPVDSPVLGQRYFLP 572

H83037
urease alpha subunit PA4868 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 06-Jan-2003
C:Accession: H83037
R:Stover, C.K.; Yam, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:2043737; PMID:10984043
A:Accession: H83037
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-566 <STO>
A:Cross-references: UNIPARC:UPI00000C5E4A; GB:AE004900; GB:AE004091; NID:g9951134; PIDN:A
A:Experimental source: strain PA01
C:Genetics:
A:Gene: ureC; PA4868
C:Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 61.3%; Score 1838; DB 2; Length 566;
Best Local Similarity 59.1%; Pred. No. 1.9e-35;
Matches 337; Conservative 88; Mismatches 139; Indels 6; Gaps 3;

QY 1 MKKKQBYVNTYGTGKGVRLGDTDLMAEVEHDYTYTGELKFGAGKTIREGQGQNSP 60
DB 1 MKISQAYAMFGPTVGDRVRLADTLWIEVERDFTVYGEVFGGKVRIDGCMGQSQ 60

QY 61 DENTLDLVITNMIIDYGIYKADIGIKNGKIHGKAGNKMDDGVSPhMVVGVTAL 120
DB 61 AAQVVDVTITNALIDHWGIVKADIGLKDGRIAIGKAGNPDMPQGV--NIAIGAGTEVI 118

QY 121 AGEPMITAGGIDSHTFSLPQQPPTALANGVTVMFGGCTGPDVGTNATITTPGKNLHR 180
DB 119 AGEMLITAGGIDSHTFSLPQQPPTALANGVTVMFGGCTGPDVGTNATITTPGKNLHR 178

QY 181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
DB 179 MLQAADAFPMNIGFTGKGNASLPPLBEQVLAGAIGLKLHEDWGTTPSAIDHCLSLVAERH 238

QY 241 DVQVCHTDTVNEAGYVDDTLNAMGRAIHAYHIEGAGGHSPPVITWAGELNLPSSST 300
DB 239 DIQVAIHDTLNEGSFVETTLGAFKGTIRHTYHTEGAGGHPADIIKACGFANVLPSTN 298

QY 301 PTIPYNTINTVAEHLDMLTCHHLDKRIREDLPQSRIKPGSIAAEDVLHDMGVIAVTSS 360
DB 299 PTRPFTNTIDEHLDMVCHHLDPAEDVAFESIRRETIAEDILHDLGAFMISS 358

QY 361 DSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
DB 359 DSQAMGRVGEVIRPTWQTADKNKQGRGLDGDGARNDFRARRVIAKYITINPAITHGISH 418

QY 421 YIGSVBEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMPGH 480
DB 419 EVGSVEAGKWADLVLRPAFFGVKPSILKGAIAAASLMGDINGSIPTPQPVHYRPMFAS 478

QY 481 HGKAKFDTSTITFVSKVAYENGVEKLGRLERQVLPVKNCRNITKKDFKNDKTAKITVDPK 540
DB 479 YAGSRATSLTTFVSOAFAAGVFPQGLRKAIGVSGRCGVQKTDLHNGLYLPTIEVDAQ 538

QY 541 TPEVFVDGKL--CTSKPTSOVPLAQRVTF 568
DB 539 NYQVRADGQLLWC--EPADVLPMAQRVFLP 566

RESULT 12
S42607
urease (EC 3.5.1.5) - Rhizobium meliloti
C:Species: Rhizobium meliloti
C>Date: 25-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S42607
R:Mikech, G.; Arnold, W.; Lentzsch, P.; Priefer, U.B.; Puhler, A.

Mol. Gen. Genet. 242, 539-550, 1994
A>Title: A 4.6 kb DNA region of Rhizobium meliloti involved in determining urease and hycing frames.
A:Reference number: S42601; MUID:94166766; PMID:8121412
A:Accession: S42607
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-570 <MIK>
A:Cross-references: UNIPROT:P42885; UNIPARC:UPI000016FFCB; GB:S69145; NID:g545795; PIDN:A
C:Genetics:
A:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase; metalloprotein; nickel
F:5-53/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted

Query Match 61.3%; Score 1838; DB 2; Length 570;
Best Local Similarity 59.8%; Pred. No. 2e-35;
Matches 340; Conservative 95; Mismatches 130; Indels 4; Gaps 3;

QY 2 MKKKQBYVNTYGTGKGVRLGDTDLMAEVEHDYTYTGELKFGAGKTIREGQGQNSPD 61
DB 4 RMSRAAYAMFGPTVGDKVRLADTLFIEVKDFTTHGEVFGGKVRIDGCMGQSQVTR 63

QY 62 E-NTLDLVITNMIIDYGIYKADIGIKNGKIHGKAGNKMDDGVSPhMVVGVTAL 120
DB 64 EGGAVDVTITNALIDHWGIVKADIGLKDGRIAIGKAGNPDMPQGV--IIVGPGTEVI 121

QY 121 AGEPMITAGGIDSHTFSLPQQPPTALANGVTVMFGGCTGPDVGTNATITTPGKNLHR 180
DB 122 AGEKIVTAGMDSHIFICPQIIEALMSGLTCLMGGTGPAHGTATCTCTGPHIAR 181

QY 181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
DB 182 MIERADA FPMNLA PAGKGNASLPGLAVEMVLGATSLKLHEDWGTTPAAIDCCLSVADEY 241

QY 241 DVQVCHTDTVNEAGYVDDTLNAMGRAIHAYHIEGAGGHSPPVITWAGELNLPSSST 300
DB 242 DVQVMIHTDLNESGFVEDTIAIKGRTHAYHTEGAGGHPADIIIRICQPNVIPSSTN 301

QY 301 PTIPYNTINTVAEHLDMLTCHHLDKRIREDLPQSRIKPGSIAAEDVLHDMGVIAVTSS 360
DB 302 PTRPVTNTLAEHLDMVCHHLSPTIPEDIAFASIRRETIAEDILHDLGAFSISS 361

QY 361 DSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
DB 362 DSQAMGRVGEVIRPTWQTADKNKQGRGLKBEETGNDNFVRKRYIAKYITINPAIAGLSH 421

QY 421 YIGSVBEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMPGH 480
DB 422 EIGSLVGRADLVLRNPAFFGVKPDVLLGGTIAAAPMGDPNASIPTPQPVHYRPMFGA 481

QY 481 HGKAKFDTSTITFVSKVAYENGVEKLGRLERQVLPVKNCRNITKKDFKNDKTAKITVDP 539
DB 482 YGRSRTNSSFVFPASLDAGLAGRLGAKELAVQNTRGGIGKASMIHNSLTPHIEVDP 541

QY 540 KTEVFVDGKLCTSKPTSOVPLAQRVTF 568
DB 542 ETYEVRADGELLCTCEPATVLPMAQRVFLP 570

RESULT 13
C36138
urease (EC 3.5.1.5) 62K chain [validated] - Klebsiella pneumoniae
N:Alternate names: urea amidohydrolase; urease alpha chain; urease chain C
C:Species: Klebsiella pneumoniae
C>Date: 30-Nov-1990 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004
C:Accession: C36138
R:Mulrooney, S.B.; Hausinger, R.P.
J. Bacteriol. 172, 5837-5843, 1990
A>Title: Sequence of the Klebsiella aerogenes urease genes and evidence for accessory protein A
A:Reference number: A36138; MUID:91008957; PMID:2211515
A:Accession: C36138

Db	539	QTVIVKADGVPLVC--EPATELPMARVPLF 567	Db	539	QTVIVKADGVPLVC--EPATELPMARVPLF 567
RESULT 15					
S08480					
urease (EC 3.5.1.5) 62K chain - Proteus vulgaris					
N:Alternate names: urease alpha chain					
C:Species: Proteus vulgaris					
C>Date: 29-Jan-1993 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004					
C:Accession: S08480					
R:Moersdorf, G.; Kaltwasser, H.					
FEMS Microbiol. Lett. 66, 67-74, 1991					
A:Title: Cloning of the genes encoding urease from Proteus vulgaris and sequencing of th					
A:Reference number: S08478; MUID:92038908; PMID:1936938					
A:Accession: S08480					
A:Molecule type: DNA					
A:Residues: 1-567 <MOE>					
A:Cross-references: UNIPROT:P16122; UNIPARC:UPI0000137D75; EMBL:X51816; NID:G45933; PIDN					
C:Genetics:					
A:Gene: ureC					
C:Complex: trimer of heterotrimeric subunits consisting of 11K (gamma), 12K (beta), and					
C:Function:					
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and tw					
C:Superfamily: urease, alpha subunit; urease 62K chain homology					
C:Keywords: heterotrimer; hydrolase; metalloprotein; nickel					
F:4-550/Domain: urease 62K chain homology <U62>					
F:134,136,217,360/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted					
F:217,246,272/Binding site: nickel 1 (Lys, His, His) #status predicted					
F:217/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predict					
F:219,320/Active site: His #status predicted					
Query Match 60.8%; Score 1822; DB 1; Length 567;					
Best Local Similarity 58.8%; Pred. No. 4.8e-35;					
Matches 336; Conservative 95; Mismatches 133; Indels 7; Gaps 4;					
Qy	1	MK-MKKQEVNTYPTKDKYRLGDTDLWAEVHDYTYGEELEKFGAGKTIREGMGQNS 59			
Db	1	MKTISRQAYADMFGPTTGDRLRLADTELEFLEIQDFTTYGEEVFGGKVRDGMGQSQV 60			
Qy	60	PDENTLDLVITNMIIDYTVIKADIGIKNGKIGIKGAKNKMDDGVSPHMVVGVTGEA 119			
Db	61	VSACVDVLIITNIIIDHWGIVKADIGIKGRITGIGKAGNPDVQPNVD--IVIGPGTEV 118			
Qy	120	LAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGTNATTITPKKWLH 179			
Db	119	VAGEGKIITAGGVDTIHIFICPQAAEEGLISGVITFIGGTPGVAGTNATTITPGIWNH 178			
Qy	180	RMLRAAEYSNNVGFGLKGNSSSKQLVQVEAGAIGPKLHEDWGTTPSAIDHCLSVADE 239			
Db	179	RMLEAVDELPIINVGLFGKVCVQPEAREIQEAGAIGLKIHEWDGATPMAIHNCINVADE 238			
Qy	240	YDVQVCIHTDTVNEAGYVDDTLNANMGRAIHAYHIEGAGGGHSPDVITMAGELNLPST 299			
Db	239	MDVQVAIHSPTLNEGGFYETVKAIAGRVIHVFHTEGAGGGHAPDVIKSVGEPNLPAST 298			
Qy	300	TPPTPYTINTVAEHLDMLTCKRDLQFQSQRIRPGSIAAEDVLHDMGVIAMTS 359			
Db	299	NPTMPYTINTVDEHDLMLVCHLDPSIPEDVFAESRIRRETIAAEDILHDMGAI SVM 358			
Qy	360	SDSQAMGRAGEVI PRTWQTADKNKFKGLPEDGKNDNFRIKRYISKYITINPALTHGV 419			
Db	359	SDSQAMGRVGEVVRWTCQCAHKMLQKGLAGDTAENDNNEIKRYIAKYITINPALH 418			
Qy	420	EYIGSVBEGKIADLVVNPAPFGVKPKIVKGMVVFSEMGDSNASVPTPOPVTYREMF 479			
Db	419	HEVGSIEKGLADIVLDPAPFFGVKPKALIMKGMVAVAPMGDINAAIPTQPVHYRPM 478			
Qy	480	HGKAKFDTSITTFVSKVAYENGVEKGLERQVLPVKNCRNITKKDKFNDKTAKITVDP 539			
Db	479	CLGRAKYQTSIMIFMSKAGIDAGVPEKGLQLGLIGRSGCRKVTKASIMHNSYVPHIE 538			
Qy	540	KTFEVFDG--KLCTSKPTSQVPLAQRVTF 568			

Search completed: August 9, 2006, 21:43:46
Job time : 43 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 21:34:37 ; Search time 305 Seconds

(without alignments)

1722.652 Million cell updates/sec

Title: US-09-904-994B-3

Perfect score: 2999

Sequence: 1 MKMKQEVYNTYGTGDKV.....KLCTSKPTSQVPLAQRVTF 568

Scoring table: BLOSUM62

Gapop 4.0 , Gapext 1.0

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	77.6	569	2 Q8GH97_HELBI	Q8gh97 helicobacte
2	2304	76.8	568	1 URE2_HELHE	P42823 helicobacte
3	2301	76.7	569	2 Q7X3W5_HELPHY	Q7x3w5 helicobacte
4	2297	76.6	569	1 URE2_HELPH	P69997 helicobacte
5	2297	76.6	569	1 URE2_HELPH	P69996 helicobacte
6	2297	76.6	569	2 Q9S0Q5_HELPHY	Q9s0q5 helicobacte
7	2293	76.5	568	2 Q38Q64_HELPHY	Q38q64 helicobacte
8	2292	76.4	567	2 Q64EY3_HELPHY	Q64ey3 helicobacte
9	2291	76.4	569	2 Q9AFB1_HELPHY	Q9afb1 helicobacte
10	2288	76.3	569	2 Q84F75_HELPHY	Q84f75 helicobacte
11	2251	75.1	559	2 Q8RNU6_HELPHY	Q8rnu6 helicobacte
12	2244	74.8	569	1 URE2_HELFE	Q87116 helicobacte
13	2109	70.3	569	2 Q93FU4_HELHP	Q93pj4 helicobacte
14	2044	68.2	569	1 URE1_BACSB	Q07397 bacillus sp
15	2027	67.6	571	2 Q9KG59_BACHD	Q9kg59 bacillus ha
16	1999	66.7	569	2 Q5KYM1_GEOKA	Q5kym1 geobacillus
17	1998	66.6	569	1 URE1_BACSU	P77837 bacillus su
18	1986	66.2	572	2 Q4CHE3_CLOTF	Q4che3 clostridium
19	1971	65.7	569	2 Q52305_SYNEP2	Q52305 synechococc
20	1942	64.8	569	1 URE1_SYNY3	P73061 synechocyst
21	1942	64.8	570	2 Q2ZGT0_CALSA	Q2zgt0 caldicellul
22	1932	64.4	568	2 Q3M712_ANAVT	Q3m712 anabaena va
23	1932	64.4	570	2 Q733J6_BACCL	Q733j6 bacillus ce
24	1930	64.4	572	2 Q8DMV6_SYNEL	Q8dmv6 synechococc
25	1929	64.3	568	2 Q3FBA4_9BURK	Q3fba4 burkholderi
26	1926	64.2	568	2 Q454M1_9BURK	Q454m1 burkholderi
27	1926	64.2	568	2 Q4LNH4_9BURK	Q4lnh4 burkholderi
28	1924	64.2	565	2 Q5FB23_CAMLA	Q5fb23 campylobact
29	1921	64.1	566	2 Q48DE6_PSEI4	Q48de6 pseudomonas
30	1921	64.1	566	2 Q87VP0_PSESM	Q87vp0 pseudomonas
31	1920	64.0	566	2 Q4ZN06_PSEU2	Q4zn06 pseudomonas

32	1920	64.0	568	2 Q39TW3_BURS3	Q39tw3 burkholderi
33	1916	63.9	568	2 Q8YQZ0_ANASP	Q8yqz0 anabaena sp
34	1914	63.8	569	2 Q46IY3_PROMT	Q46iy3 prochloroco
35	1910	63.7	568	2 Q4BIH6_BURVI	Q4bih6 burkholderi
36	1905	63.5	566	2 Q3KIT2_PSBPF	Q3kit2 pseudomonas
37	1905	63.5	568	2 Q34XU6_9GAMM	Q34xu6 alkalilimni
38	1905	63.5	568	2 Q63RL3_BURPS	Q63rl3 burkholderi
39	1904	63.5	572	2 Q2JQ88_9CVAN	Q2jq88 cyancobacter
40	1902	63.4	566	2 Q4KJ10_PSEPF	Q4kj10 pseudomonas
41	1902	63.4	568	2 Q6ZHS0_BURMA	Q6zhs0 burkholderi
42	1901	63.4	564	2 Q3JPU6_BURP1	Q3jpu6 burkholderi
43	1900	63.4	566	2 Q4IXD2_AZOV1	Q4ixd2 azotobacter
44	1900	63.4	569	1 URE1_BACPA	P41020 bacillus pa
45	1891	63.1	570	2 Q3WT52_9RHIZ	Q3wt52 mesorhizobi

ALIGNMENTS

RESULT 1
Q8GH97_HELBI PRELIMINARY; PRT; 569 AA.

AC Q8GH97;

DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.

DT 01-MAR-2003, sequence version 1.

DT 07-FEB-2006, entry version 14.

DE Urease B.

OS Helicobacter bizzozeronii.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC Helicobacteraceae; Helicobacter.

OX NCBI_TaxID=56877;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22540228; PubMed=12652903; DOI=10.1080/10425170210000392230;

RA Zhu J., Teng C.H., Chang C.F., Chang C.D., Simpson K.W., Wei C.,

RA McDonough P., McDonough S., Chang Y.F.

RT "Cloning and characterization of a Helicobacter bizzozeronii urease

gene cluster.";

RL DNA Seq. 13:321-331(2002).

CC Copyright by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; AF330621; AA015374.1; -; Genomic_DNA.

DR HSSP; P14917; IE9Y.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0016151; F:nickel ion binding; IEA.

DR GO; GO:0009039; E:urease activity; IEA.

DR GO; GO:0019627; P:urea metabolism; IEA.

DR InterPro; IPR006680; Amidohydro 1.

DR InterPro; IPR011612; Urease_alpha N.

DR InterPro; IPR008295; Urease_alphaIone.

DR Pfam; PF01979; Amidohydro 1.

DR Pfam; PF00449; Urease_alpha; 1.

DR PIRSF; PIRSF001226; Urease_alpha; 1.

DR PRINTS; PR01752; UREASE.

DR TIGRFAMs; TIGR01792; urease_alph; 1.

DR PROSITE; PS01120; UREASE_1; 1.

DR PROSITE; PS00145; UREASE_2; 1.

SQ SEQUENCE 569 AA; 61922 MW; 3938FFC83EC39E26 CRC64;

Query Match 77.6%; Score 2326; DB 2; Length 569;

Best Local Similarity 74.1%; Pred No. 3.6e-43;

Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;

Qy 2 MKMKQEVYNTYGTGDKVRLGDTDLAEVEHDYTYGELKFGAGKTTIREGMQNSPD 61

3 KISRKEYVMYGTGDKVRLGDTDLILEVHDCTTYGEEKFGGKTTIRDGMAQTNSPS 62

62 ENTLDLVTNMIIDYGIYKADIGIKNGKIHGICAGKMDQGVSPHMVVGVTGTEALA 121

63 SHELDVLVTNALIVDYGTYGIYKADIGIKNGKIHGICAGKMDQGVSPHMVVGVTGTEALA 122

```

Qy 122 GEGMIITAGGIDSHFLSPQOFTALANGVTTWFGGTPGVDGTNATTITPGKWNLRHM 181
Db 123 AEGLIIVTAGGIDTHIFISPOQITAFASGITIMGGTGPADGTNATTITPGRWNLRKM 182
Qy 182 LRAAEYSMNVGFLGKGNSSKKOLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADDEYD 241
Db 183 LRAEEYAMNLGYLGKGNVSEPSLVQLEAGAIGFKLHEDWGTTPAAIYHCLNVADKID 242
Qy 242 VQVCIHDTVNEAGYVDDTLNMMNGRAIHAYHIEGAGGSHSPDVTWAGELNLPSTTP 301
Db 243 VQVAIHTDLNEACGVEDTLQAIAGRTIHTHTTEGAGGSHADPVIKMSGEFNILPASTNP 302
Qy 302 TTPYTNVVAEHLDMTCHLDKRIREDLPQSQRIRPGSIAEDVLHDMGVAMTSSD 361
Db 303 TTPYTNVTEAHEMDLVMCHLDKRIREDLPQSQRIRPGSIAEDVLHDMGVAMTSSD 362
Qy 362 SQAMGRAGEVTPRTWQTADKNNKKEFGKLPEDGKNDNFRIKRYISKYTNINPALTHGVSEY 421
Db 363 SQAMGRAGEVTPRTWQTADKNNKKEFGKLPEDGKNDNFRIKRYISKYTNINPALTHGVSEY 422
Qy 422 IGSVEEGKIADLVVNNPAFFGKPKIVIKGGMVFSEMGSDSNASVPTPQPVYVREMFGHH 481
Db 423 VGSVEVGKPADLVLSAPFGIKENMIKGGFIALSQMGDANASIPTPQPVYVREMFGHH 482
Qy 482 GKAKFDTSTITFVSKVAYENGVEKLGLEQVLPVKNCRNITTKDFKNDKTAKITVDPKT 541
Db 483 GKAKFDTNITFVSKVAYENGVEKLGLEQVLPVKNCRNITTKDFKNDKTAKITVDPKT 542
Qy 542 FEVFDGKLCITSKPTSQVPLAQRTFF 568
Db 543 YKVKVDGKEVTSKAADKISLAQLNLF 569

RESULT 2
URE2_HELHE
ID_URE2_HELHE STANDARD; PRT; 568 AA.
AC P42823;
DT 01-NOV-1995, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1995, sequence version 1.
DT 07-FEB-2006, entry version 44.
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureb;
OS Helicobacter heilmannii.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=35817;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=2;
RC MEDLINE=94222523; PubMed=8168924;
RX Solnick J.V., O'Rourke J., Lee A., Tompkins L.S.;
RA "Molecular analysis of urease genes from a newly identified uncultured
RT species of Helicobacter.";
RL Infect. Immun. 62:1631-1638 (1994).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COPACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: Consists of two subunits (alpha and beta).
CC -!- FM: Lys-219 is carbamylated. The carbamoyl group provides the
CC ligands for the two nickel ions (By similarity).
CC -!- SIMILARITY: Belongs to the urease family.
CC -!- CAUTION: In Helicobacter the beta subunit is what is known in
CC other bacteria as the alpha subunit.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: L25079; AAA65723.1; -; Genomic_DNA.
CC HSSP: P14917; 1E9Z.
CC InterPro: IPR006680; Amidohydro 1.
CC InterPro: IPR005848; Urease_alpha.
CC InterPro: IPR011612; Urease_alpha_N.
CC InterPro: IPR008295; Urease_alphalone.

```

```

DR Pfam: PF01979; Amidohydro 1; 1.
DR Pfam: PF00449; Urease_alpha; 1.
DR PIRSF: PIRSF001226; Urease_alpha; 1.
DR PRINTS: PR01752; UREASE.
DR TIGRFAMs: TIGR01792; urease_alpha; 1.
DR PROSITE: PS01120; UREASE_1; 1.
DR PROSITE: PS00145; UREASE_2; 1.
KW Hydrolase; Metal-binding; Nickel.
FT CHAIN 1 568 Urease beta subunit.
FT ACT_SITE 321 321 Proton donor (By similarity).
FT METAL 136 136 Nickel 2 (By similarity).
FT METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 (By similarity).
FT METAL 219 219 Nickel 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 361 361 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
SQ SEQUENCE 568 AA; 61870 MW; E2C036DF30377917 CRC64;

Query Match 76.8%; Score 2304; DB 1; Length 568;
Best Local Similarity 73.7%; Pred. No. 1.1e-42;
Matches 418; Conservative 74; Mismatches 74; Indels 1; Gaps 1;

Qy 2 KMKQDEYVNTYGTGDKVRLGDTDLMAVEHHDYTYTGELKPGAGKTIREGQNSPD 61
Db 3 KISKEYVSMYGTPTGDKVRLGDTDLMAVEHHDYTYTGELKPGAGKTIREGQNSPD 62
Qy 62 ENTLDLVITNAMIIDYTYGKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEALA 121
Db 63 SHELDLVITNALIVDYTYGIKADIGIKNGKIHGIGKAGNKLDDGVNCLRCVGPTEALA 122
Qy 122 GEGMIITAGGIDSHFLSPQOFTALANGVTTWFGGTPGVDGTNATTITPGKWNLRHM 181
Db 123 AEGLIIVTAGGIDTHIFISPOQITAFASGITIMGGTGPADGTNATTITPGRWNLRKM 182
Qy 182 LRAAEYSMNVGFLGKGNSSKKOLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADDEYD 241
Db 183 LRAEEYAMNLGYLGKGNVSEPSLVQLEAGAIGFKLHEDWGTTPSAIDHCLSVADDEYD 242
Qy 242 VQVCIHDTVNEAGYVDDTLNMMNGRAIHAYHIEGAGGSHSPDVTWAGELNLPSTTP 301
Db 243 VQVAIHTDLNEACGVEDTLQAIAGRTIHTHTTEGAGGSHADPVIKMSGEFNILPASTNP 302
Qy 302 TTPYTNVVAEHLDMTCHLDKRIREDLPQSQRIRPGSIAEDVLHDMGVAMTSSD 361
Db 303 TTPYTNVTEAHEMDLVMCHLDKRIREDLPQSQRIRPGSIAEDVLHDMGVAMTSSD 362
Qy 362 SQAMGRAGEVTPRTWQTADKNNKKEFGKLPEDGKNDNFRIKRYISKYTNINPALTHGVSEY 421
Db 363 SQAMGRAGEVTPRTWQTADKNNKKEFGKLPEDGKNDNFRIKRYISKYTNINPALTHGVSEY 422
Qy 422 IGSVEEGKIADLVVNNPAFFGKPKIVIKGGMVFSEMGSDSNASVPTPQPVYVREMFGHH 481
Db 423 VGSVEVGKPADLVLSAPFGIKENMIKGGFIALSQMGDANASIPTPQPVYVREMFGHH 482
Qy 482 GKAKFDTSTITFVSKVAYENGVEKLGLEQVLPVKNCRNITTKDFKNDKTAKITVDPKT 541
Db 483 GKAKFDTNITFVSKVAYENGVEKLGLEQVLPVKNCRNITTKDFKNDKTAKITVDPKT 542
Qy 542 FEVFDGKLCITSKPTSQVPLAQRTFF 568
Db 543 YKVKVDGKEVTSKAADKISLAQLNLF 568

RESULT 3
Q7X3W5_HELPY
ID Q7X3W5_HELPY PRELIMINARY; PRT; 569 AA.
AC Q7X3W5;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.

```

```
DE Urease B.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MEL-HP27;
RA Duan G.C., Dai L.P.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY295085; AAP51176.1; -; Genomic_DNA.
DR HSSP; P14917; 189Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016351; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; F:urea metabolism; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha_N.
DR InterPro; IPR008295; Urease_alpha1one.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
DR SEQUENCE 569 AA; 61671 MW; 4C96852826C7A784 CRC64;
Query Match 76.7%; Score 2301; DB 2; Length 569;
Best Local Similarity 73.5%; Pred No. 1.3e-42;
Matches 417; Conservative 75; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVHDYTTYGEELKFGAGKTIREGMQSNPSD 61
DB 3 KISKEVSVSYGPTGDKVRLGDTDLWAEVHDYTTYGEELKFGAGKTIREGMQSNPS 62
QY 62 ENTLDLVITNAMIIDYTGIIYKADIGIKNGIKHGKAGNKMQDGVSPHMVVGVTAL 121
DB 63 KEELDLIITNALIVDYTGIIYKADIGIKNGIKAGTIGKGNKMDQGVKNLSVGPATE 122
QY 122 GEGMILITAGGIDSHTHFELSPQEPPTALANGVTYTMFGGGTGPVDCGTNATTTPG 181
DB 123 GEGILVTAGGIDTHIHPISPQIPTAFASGVTTMIGGGTGPADGTNATTTPGRRNL 182
QY 182 LRAAEYSNMVGLFGKGNSSKKQLVQVEAGAGFKLHEDWGTTPSAIDHCLSVADBYD 241
DB 183 LRAAEYSNMVGLFLAGNANDASLADQIEAGALGFKIHEDWGTTPSAINHLDVADYD 242
QY 242 VQVCIIHTDVNEAGYVDDTLNANMGRAIHAYHIEGAGGSHSPDVITWAGELNLPSS 301
DB 243 VQVAIHTDLNEAGVEDTMAAIGRTHMTFHTEGAGGCHAPDIIVKAGEHNILPASTNP 302
QY 302 TIPTTINTVAEHLDMLTCHLLOKRRREDLQFSQSRIRPGSIAAEVDLHMGVJAMTSSD 361
DB 303 TIPTVTNTEAHEMDMLVMCHLHDKSIEDVQPADSIRPQTIAEDTLHDMGIFSISSD 362
QY 362 SQAMGRAGEVIRPTWQTADKNKKEFGKLPELGDGNDNFRIKRYISKYITINPALTHGV 421
DB 363 SQAMRGVEVIRPTWQTADKNKKEFGKLKEKGDNDNFRIKRYLSKYITINPALTHGVSE 422
QY 422 IGSVEEGKIADLVNMPAPFGVKKPKIVKGMVVFSEMGDSNASVPTPQPVYVREMGHH 481
DB 423 VGSVEGVKADLVLMSPAPFGVKKPKIVKGMVVFSEMGDSNASVPTPQPVYVREMPAHH 482
QY 482 GKAKFTDSTTFVSKVAYENGVKELGLERQVLVPKNCRNITKKDFKNDKTAKITVDPKT 541
DB 483 GKAKYDANITFVSKAAAYDKIGBELGLERQVLVPKNCRNITKKDMQFNDTTHAIEVNPET 542
```

```
QY 542 FEFVFDGKLGKTSKPTSQVPLAQRYTTP 568
DB 543 YHVFVDGKGVTSKPTKVKSLAQLFISIP 569
RESULT 4
URE2_HELPJ STANDARD; PRT; 569 AA.
ID URE2_HELPJ
AC P69997; P14917;
DT 04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT 04-JAN-2005, sequence version 1.
DT 07-MAR-2006, entry version 7.
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureB; Synonyms=hpb; OrderedLocNames=JHP0067;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nichelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: Heterohexamer of 3 alpha and 3 beta subunits. Four
CC heterohexamers assemble to form a 16 nm spherical complex (By
CC similarity).
CC -!- PTM: Lys-219 is carbamylated. The carbamoyl group provides the
CC ligands for the two nickel ions (By similarity).
CC -!- MISCELLANEOUS: Oligomerization may protect the enzyme against
CC denaturation in an acidic environment (By similarity).
CC -!- SIMILARITY: Belongs to the urease family.
CC -!- CAUTION: In Helicobacter the beta subunit is what is known in
CC other bacteria as the alpha subunit.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AE001446; AAD05651.1; -; Genomic_DNA.
PIR; B38537; URKCBP.
GenomeReviews; AE001439 GR; JHP0067.
InterPro; IPR006680; Amidohydro_1.
InterPro; IPR005848; Urease_alpha.
InterPro; IPR011612; Urease_alpha_N.
InterPro; IPR008295; Urease_alpha1one.
Pfam; PF01979; Amidohydro_1; 1.
Pfam; PF00449; Urease_alpha; 1.
PIRSF; PIRSF001226; Urease_alpha; 1.
PRINTS; PR01752; UREASE.
TIGRFAMs; TIGR01792; urease_alph; 1.
PROSITE; PS01120; UREASE_1; 1.
PROSITE; PS00145; UREASE_2; 1.
KW Complete proteome; Hydrolase; Metal-binding; Nickel.
CHAIN 1 569
Urease beta subunit.
FT ACT_SITE 322 322 Proton donor (By similarity).
FT METAL 136 136 Nickel 2 (By similarity).
FT METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 (By similarity).
FT METAL 219 219 Nickel 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 362 362 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
SQ SEQUENCE 569 AA; 61684 MW; 4C8A6BC6C8295584 CRC64;
```

[illegible]

RESULT 5

```

URE2_HELPY
ID      URE2_HELPY      STANDARD;      PRT;      569 AA.
AC      P69996; P14917; Q9R3B3;
DT      04-JAN-2005, integrated into UniProtKB/Swiss-Prot.
DT      04-JAN-2005, sequence version 1.
DT      07-MAR-2006, entry version 11.
DE      Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN      Name=ureB; Synonyms=hpb; OrderedLocusNames=HP0072;
OS      Helicobacter pylori (Campylobacter pylori).
OC      Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC      Helicobacteraceae; Helicobacter.
OX      NCBI_TaxId=210;
RX      [1]
RN      NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RP      STRAIN=85P;
RC      MEDLINE=91161505; PubMed=2001995;
RX      Labigne A., Cussac V., Courcoux P.;
RT      "Shuttle cloning and nucleotide sequences of Helicobacter pylori genes
RT      responsible for urease activity.";
RL      J. Bacteriol. 173:1920-1931(1991).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC      STRAIN=CPM630;
RX      MEDLINE=90221820; PubMed=2326167;
RA      Clayton C.L., Pallen M.U., Kleanthous H., Wren B.W., Tabaqchali S.;
RT      "Nucleotide sequence of two genes from Helicobacter pylori encoding

```

for urease subunits.";

RL Nucleic Acids Res. 18:362-362(1990).

RL [3]

RL NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=26695 / ATCC 700392;

RX MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;

RA Tomb J.-F., White O., Kervilave A.R., Clayton R.A., Sutton G.G.,

RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,

RA Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,

RA Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,

RA Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,

RA Hickey E.K., Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D.,

RA Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,

RA Watthey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,

RA Smith H.O., Fraser C.M., Venter J.C.;

RT "The complete genome sequence of the gastric pathogen *Helicobacter pylori*."

RL Nature 388:539-547(1997).

RL [4]

RP PROTEIN SEQUENCE OF 1-30, AND KINETIC PARAMETERS.

RP PubMed=1857197; DOI=10.1016/0882-4010(91)90062-F;

RX Evans D.J. Jr., Evans D.G., Kirkpatrick S.S., Graham D.Y.;

RA "Characterization of the *Helicobacter pylori* urease and purification of its subunits."

RT of its subunits."

RL Microb. Pathog. 10:15-26(1991).

RL [5]

RP PROTEIN SEQUENCE OF 1-20.

RP MEDLINE=90264448; PubMed=2188975;

RX Dunn B.E., Campbell G.P., Perez-Perez G.I., Blaser M.J.;

RA "Purification and characterization of urease from *Helicobacter pylori*."

RT J. Biol. Chem. 265:9464-9469(1990).

RL [6]

RP PROTEIN SEQUENCE OF 1-15.

RP MEDLINE=90202165; PubMed=2318539;

RX Hu L.T., Mobley H.L.T.;

RA "Purification and N-terminal analysis of urease from *Helicobacter pylori*."

RT Infect. Immun. 58:992-998(1990).

RL [7]

RP PROTEIN SEQUENCE OF 1-12.

RC STRAIN=NCTC 11637;

RX PubMed=1452359;

RA Turbett G.R., Hoj P.B., Horne R., Mee B.J.;

RT "Purification and characterization of the urease enzymes of *Helicobacter* species from humans and animals."

RL Infect. Immun. 60:5259-5266(1992).

RL [8]

RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 565-569.

RC STRAIN=85P;

RC MEDLINE=92210488; PubMed=1313413;

RA Cussac V., Ferrero R.L., Labigne A.;

RT "Expression of *Helicobacter pylori* urease genes in *Escherichia coli* grown under nitrogen-limiting conditions."

RL J. Bacteriol. 174:2466-2473(1992).

RL [9]

RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS), AND SUBUNIT STRUCTURE.

RP MEDLINE=121267477; PubMed=11373617; DOI=10.1038/88563;

RA Ha N.-C., Oh S.-T., Sung J.Y., Cha K.A., Lee M.H., Oh B.-H.;

RT "Supramolecular assembly and acid resistance of *Helicobacter pylori* urease."

RL Nat. Struct. Biol. 8:505-509(2001).

CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).

CC -1- COPACITOR: Binds 2 nickel ions per subunit.

CC -1- BIOPHYSICOCHEMICAL PROPERTIES:

CC Kinetic parameters:

CC KM=0.48 mM for urea;

CC pH dependence:

CC Optimum pH is 8.0. Active from pH 4.0 to 10.0;

CC -1- SUBUNIT: Heterohexamer of 3 alpha and 3 beta subunits. Four heterohexamers assemble to form a 16 nm spherical complex.

CC -1- PM: Lys-219 is carbamylated. The carbamoyl group provides the ligands for the two nickel ions (by similarity).


```

RESULT 6
Q9SQ05 HELPY PRELIMINARY; PRT; 569 AA.
AC Q9SQ05
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Urease B.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=HPK5;
RC MEDLINE=20305050; PubMed=10844692;
RA Akada J.K., Shirai M., Takeuchi H., Tsuda M., Nakazawa T.;
RT "Identification of the urease operon in Helicobacter pylori and its
RT control by mRNA decay in response to pH.";
RL Mol. Microbiol. 36:1071-1084 (2000).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC EMBL; AB032429; BAA04533.1; -; Genomic_DNA.
DR HSSP; P14917; 189Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
DR InterPro; IPR006680; Amidohydro 1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha N.
DR InterPro; IPR008295; Urease_alpha1.
DR Pfam; PF01979; Amidohydro 1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF01226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRfams; TIGR01792; urease_alpha; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 569 AA; 61671 MW; 4C8A6BDA3A295584 CRC64;

Query Match 76.6%; Score 2297; DB 2; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.6e-42;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

Qy 2 KMKKQEVNTYGTGDKVRLGDTDLWAERVDHYTYGELKFGAGKTIREGMGOSNSPD 61
Db 3 KISRKEYSMYGTGDKVRLGDTDLIAEVEHDYTYGELKFGGKTIREGMSQSNPS 62
Qy 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTAL 121
Db 63 KEELDLIITNALIVDYGTYGIYKADIGIKGKIAGIKGGNKMDDGVKNLSVGPATEALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQPFPTALANGVTMTFGGGTGPVDGNTATTITPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTTMTGGGTGPADGTNATTITPGRNLKWM 182
Qy 182 LRAAEYSNMVGFCKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADKYD 241
Db 183 LRAAEYSNMVGFCKGNASNDASLADQIEAGAIKFKHEDWGTTPSAINHALDVADKYD 242
Qy 242 VQVCIIHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGCHSPDVTIMAGELNIPSSSTTP 301
Db 243 VQVAIHTDLNEAGCVETMAAIAAGTMTHTFTEGAGGGHAPDIIKVAGEHNILPASTNP 302
Qy 302 TIPTVINTVAEHLDMMLTCHHLDKRIREDLPQSQRIRPGSIAAEDVLHDMGVIAWTSDD 361
Db 303 TIPTVNTAEHMDMLVCHHLDKSIKEDVQPADSRIRPQTIAAEDTLHDMGIFSISSD 362

Qy 362 SQAMGRAGEVIRPTWQTADKNKKFKLPEDGKDNDFRIKRYISKYITNPALTHGVSEY 421
Db 363 SQAMGRVGEVIRPTWQTADKNKKFKLKEKGDNDNFRIKRYLSKYITNPALAHGISEY 422
Qy 422 TGSVEGKIADLVVNPAPFFGKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMGHH 481
Db 423 VGSVEVGKVDLVVNSPAPFFGKPNMIKGGFIALSQMGDANASITPQPVYVREMFHH 482
Qy 482 GKAKEDTSTTFVSKVAYENGKVKGLERQVLVKNCRNITKKDFKNDKTKATITVDPKT 541
Db 483 GKAKYDANITTFVSOAYDKIKERIGLERQVLVKNCRNITKKDMQFNDDTAHIEVNPET 542
Qy 542 FEVFDGKGLCTSKPTSQVPLAQRYTF 568
Db 543 YHVFVDGKEVTSKPKATKVSQAQLFSIF 569

RESULT 7
Q38Q64 HELPY PRELIMINARY; PRT; 568 AA.
AC Q38Q64;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Urease B (Fragment).
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=NCTC 11639;
RA Li Y., Ning Y.-S., Long M., Wang Y.-D., Dong W.-Q., Li M.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC EMBL; DQ141576; ABA87133.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
FT NON_TER 1
FT NON_TER 568
SQ SEQUENCE 568 AA; 61524 MW; 5CD9DC062D41P883 CRC64;

Query Match 76.5%; Score 2293; DB 2; Length 568;
Best Local Similarity 73.2%; Pred. No. 2e-42;
Matches 415; Conservative 76; Mismatches 76; Indels 0; Gaps 0;

Qy 2 KMKKQEVNTYGTGDKVRLGDTDLWAERVDHYTYGELKFGAGKTIREGMGOSNSPD 61
Db 2 KISRKEYSMYGTGDKVRLGDTDLIAEVEHDYTYGELKFGGKTIREGMSQSNPS 61
Qy 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTAL 121
Db 62 KEELDLIITNALIVDYGTYGIYKADIGIKGKIAGIKGGNKMDDGVKNLSVGPATEALA 121
Qy 122 GEGMIITAGGIDSHTHFLSPQPFPTALANGVTMTFGGGTGPVDGNTATTITPGKNLHRM 181
Db 122 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTTMTGGGTGPADGTNATTITPGRNLKWM 181
Qy 182 LRAAEYSNMVGFCKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADKYD 241
Db 182 LRAAEYSNMVGFCKGNASNDASLADQIEAGAIKFKHEDWGTTPSAINHALDVADKYD 241
Qy 242 VQVCIIHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGCHSPDVTIMAGELNIPSSSTTP 301
Db 242 VQVAIHTDLNEAGCVETMAAIAAGTMTHTFTEGAGGGHAPDIIKVAGEHNILPASTNP 301
Qy 302 TIPTVINTVAEHLDMMLTCHHLDKRIREDLPQSQRIRPGSIAAEDVLHDMGVIAWTSDD 361

```

```

Db 302 TIPTVNTAEHMDMLVCHHLDKSIKEDVQFADSRIRPQIIRPQIAEDTLHDMGIFSISSD 361
Qy 362 SQAMGRAGEVITPRTWQTADKNKGFGLPEBGKONDNFRIKRYISKYTYINPALTHGVSEY 421
Db 362 SQAMGRVGEVITRTWQTADKNKGFGLPEBGKONDNFRIKRYISKYTYINPAIAHGISEY 421
Qy 422 IGSVEEGKIADLVVNPAPFGVKPKIVIKGVMVFSEMGDSNASVPTPQPVYREMGHH 481
Db 422 VGSVEVGKVALDLVWSPAFFGKPNMIIKGFIALSQMGDANASITPTPQPVYREMPAHH 481
Qy 482 GKAKFDTSITFVSKVAYENGVEKGLERQVLPVKNCRNITKQDFKNDKTAKIIVDPKT 541
Db 482 GKAKYDANITFVSOAAYDKGKKEELGLERQVLPVKNCRNITKQDMQFNDTTHAIEVNPET 541
Qy 542 FEVFDGKLCSTKPTSQVPLAQRYTF 568
Db 542 YHVFVDGKEVTSKPNKVSIAQLFSIF 568

RESULT 8
Q64EY3 HELPY
ID Q64EY3 HELPY PRELIMINARY; PRT; 567 AA.
AC Q64EY3
DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE UreB.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Y06;
RX PubMed=15052678;
RA Mao Y.F., Yan J.;
RT "Construction of prokaryotic expression system of ureB gene from a
RT clinical Helicobacter pylori strain and identification of the
RT recombinant protein immunity.";
RL World J. Gastroenterol. 10:977-984 (2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; AY714324; AU21200.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
DR InterPro; IPR006680; Amidohydro 1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha N.
DR InterPro; IPR008295; Urease_alphalone.
DR Pfam; PF01979; Amidohydro 1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 567 AA; 61444 MW; 1DPAIAA596205A9C CRC64;

Query Match 76.4%; Score 2292; DB 2; Length 567;
Best Local Similarity 73.3%; Pred. No. 2,le-42;
Matches 415; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

Qy 3 MKQEVNTYPTGDKVRLGDTDLAEVHDYTYIGEELKFGAGKTIREGMGQSNDPE 62
Db 2 ISRKEYVMYGPPTGDKVRLGDTDLAEVHDYTYIGEELKFGAGKTIREGMGQSNDPSK 61
Qy 63 NTLDLVTNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMQDQGVSPHVVGVGTALAG 122

```

```

Db 62 EELDLIITNALIVDYGIIYKADIGIKNGKIHGIGKAGNKMQDQGVKNLISVGPATEALAG 121
Qy 123 EGMITTAGIDSHTHFLSPQQPTALANGVTTFMGGGTGVGDTNATITTPGKNLHRLM 182
Db 122 EGLIVTAGIDTHIHFIISPPQIPTAFASGVTTMIGGTGPDGNTATITTPGRNLKMWL 181
Qy 183 RAAEYSNMVGLFGKNSSSKKQLVEOVEACAGIPLKLEDHGWGTPSAIDHCLSVADYDV 242
Db 182 RAAEYSNMVGLFLAKGNTSNDASLADQIEAGIPLKLEDHGWGTPSAINHALDVADYDV 241
Qy 243 QVCITHTVNEAGYVDDTLNAMNGRAIHAVHIEGAGGSHSPDVTWAGELNILLPSSTPT 302
Db 242 QVAIHTDLNEAGCVEDTAAIAGRTMHTFHTEGAGGCHAPDIIKVAGENIILPASTNPT 301
Qy 303 IPTYTINTVAEHLMDLMTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSDS 362
Db 302 IFFTNTAEHMDMLVCHHLDKSIKEDVQFADSRIRPQIIRPQIAEDTLHDMGIFSISSDS 361
Qy 363 QAMGRAGEVITPRTWQTADKNKGFGLPEBGKONDNFRIKRYISKYTYINPALTHGVSEYI 422
Db 362 QAMGRVGEVITRTWQTADKNKGFGLPEBGKONDNFRIKRYISKYTYINPAIAHGISEYV 421
Qy 423 GSVEGKIADLVVNPAPFGVKPKIVIKGVMVFSEMGDSNASVPTPQPVYREMGHHG 482
Db 422 GSVEVGKVALDLVWSPAFFGKPNMIIKGFIALSQMGDANASITPTPQPVYREMPAHHG 481
Qy 483 KAKFDTSITFVSKVAYENGVEKGLERQVLPVKNCRNITKQDFKNDKTAKITVDPKTF 542
Db 482 KAKYDANITFVSOAAYDKGKKEELGLERQVLPVKNCRNITKQDMQFNDTTHAIEVNPETY 541
Qy 543 EVFVDGKLCSTKPTSQVPLAQRYTF 568
Db 542 YHVFVDGKEVTSKPNKVSIAQLFSIF 567

RESULT 9
Q9AFB1 HELPY
ID Q9AFB1 HELPY PRELIMINARY; PRT; 569 AA.
AC Q9AFB1
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Urease B.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gerold A., Simpson K., Chang Y.-F.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; AF352376; AAK32714.1; -; Genomic DNA.
DR HSSP; P14917; IE9Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
DR InterPro; IPR006680; Amidohydro 1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha N.
DR InterPro; IPR008295; Urease_alphalone.
DR Pfam; PF01979; Amidohydro 1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.

```


Db	363	SOAMRGVGEVITRTWQTADKNKKEFGRLKEBKGDNDNFRIKRYLSKYTNPAIAHGISEY	422
Qy	422	IGSVEEGKIADLVVNPAFGVQKPKIVTKGGWVVFSEMGDSNASVPTQPVTYREMFQGH	481
Db	423	VGSVEVGKADLVVNSPAFFGVQKPMIIGKGFIALSQMGDNANSLPTQPVTYREMFQGH	482
Qy	482	GKAKFTDTSITFVSQVAYENGKVKELGLERQVLPVKNCRNITKQPKFNDKTAKITVDPKT	541
Db	483	GKAKYDANIITFVSQAAYDKGKKEELGLERQVLPVKNCRNITKQMFNDTTHAIEVNPET	542
Qy	542	FEVFDGKLCIT 552	
Db	543	YHVFVDGKEVT 553	

RESULT 12

URE2_HELPFE

ID	URE2_HELPFE	STANDARD;	PRT;	569 AA.
AC	Q08716; Q9R3B3;			
DT	01-OCT-1994,	integrated into UniProtKB/Swiss-Prot.		
DT	01-OCT-1994,	sequence version 1.		
DT	07-FEB-2006,	entry version 47.		
DE	Urease beta subunit (EC 3.5.1.5)	(Urea amidohydrolase).		
GN	Name=ureB;			
OS	Helicobacter felis.			
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;			
OC	Helicobacteraceae; Helicobacter.			
OX	NCBI_TaxID=214;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [GENOMIC DNA].			
RX	MEDLINE=94018627; PubMed=8412683;			
RA	Ferrero R.L., Labigne A.;			
RT	"Cloning, expression and sequencing of Helicobacter felis urease			
RT	genes.";			
RL	Mol. Microbiol. 9:323-333(1993).			
RL	[2]			
RC	PROTEIN SEQUENCE OF 1-12.			
RC	STRAIN=ATCC 49179;			
RX	PubMed=1452359;			
RA	Turbett G.R., Hoef P.B., Horne R., Mee B.J.;			
RT	"Purification and characterization of the urease enzymes of			
RT	Helicobacter species from humans and animals.";			
CC	Infect. Immun. 60:5259-5266(1992).			
CC	-1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).			
CC	-1- COPACTOR: Binds 2 nickel ions per subunit (By similarity).			
CC	-1- SUBUNIT: Consists of two subunits (alpha and beta).			
CC	-1- PTM: Lys-219 is carbamylated. The carbamoyl group provides the			
CC	ligands for the two nickel ions (By similarity).			
CC	-1- SIMILARITY: Belongs to the urease family.			
CC	-1- CAUTION: In Helicobacter the beta subunit is what is known in			
CC	other bacteria as the alpha subunit.			

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL;	X69080; CAA48826.1; -;	Genomic_DNA.
PIR;	C49215; C49215.	
PIR;	S35291; S35291.	
HSSP;	PI4917; 1B9Z.	
InterPro;	IPR006680; Amidohydro.1.	
InterPro;	IPR005848; Urease_alpha.	
InterPro;	IPR011612; Urease_alpha_N.	
InterPro;	IPR008295; Urease_alpha1one.	
Pfam;	PF001979; Amidohydro.1; 1.	
Pfam;	PF00449; Urease_alpha; 1.	
PIRS;	PIRSF001226; Urease_alpha; 1.	
PRINTS;	PR01752; UREASE.	
TIGRFAMS;	TIGR01792; urease_alpha; 1.	
PROSITE;	PS01120; UREASE_1; 1.	
PROSITE;	PS00145; UREASE_2; 1.	

Direct protein sequencing; Hydrolase; Metal-binding; Nickel.

CHAIN 1 569 Urease beta subunit.

[illegible]

DOI=10.1128/JAI.69.9.5914-5920.2001;
Beckwith C.S., McGee D.J., Mobley H.L.T., Riley L.K.;
RT "Cloning, expression, and catalytic activity of Helicobacter hepaticus
urease.";
RL Infect. Immun. 69:5914-5920(2001).
RN [2].

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 51449 / 3B1;
RP MEDLINE=42709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA Surbaum S., Josenhans C., Stenzenbach T., Drescher B., Brandt P.,
RA Bell M., Droge M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koennig J., Machok L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AF326556; AAAG69199.1; -; Genomic_DNA.
DR ENMBL; AE017125; AAP77005.1; -; Genomic_DNA.
DR HSPP; P14917; I59Y.
DR HSPP; P18314; IASN.
DR BioCyc; RHEP235279:HHO408-MONOMER; --
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016781; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
DR InterPro; IPRO06580; Amidohydro_1.
DR InterPro; IPRO05848; Urease_alpha.
DR InterPro; IPRO11612; Urease_alpha_N.
DR InterPro; IPRO08295; Urease_alphalane.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PRSP; PRSPF001226; Urease_alpha; 1.
DR PRINTS; PRO1752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Complete proteome; Hydrolase.

SEQUENCE 569 AA; 61510 MW; 2825DIE6406F6429 CRC64;

Query Match 70.3%; Score 2109; DB 2; Length 569;
Best Local Similarity 68.1%; Pred. No. 2.8e-38;
Matches 388; Conservative 83; Mismatches 95; Indels 4; Gaps 6

Gy 1 MKMKKQEVNTYTGTGDKVRLGDTDLMAEVEHDYTTTGEELKFAGAKTIIRMGQSNSP 60
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
2 IKSRQYASMYGPTTGDKVRLGDTNLFAETKEDYTLXEEIKFGGGKTIRDGAOSAST 61

Gy 61 DENTLDIVTNAMIIDYTYGVKADIGIKNGKHGIKGKNKDMDQGVSPHMVVGVCTEAL 120
Db 62 YTNELDAVTNMIIDYTYGVKADIGIKGKLVLGKAGNPDTQDSNVEMVVGAAATEVI 121

Gy 121 AGEQMIIITAGGDISHTHFLSPQQFFPALANGVTTFMFGGSGTPVDGNATTITPGKWNLR 180
Db 122 AGEQIITAGGDITHIFISPTQIPTALYSVTTWIMGGTGPAAGTAATTCCTPGKWNMQ 181

Gy 181 MLRAAEYSMVVGLFKGNSSKKQLVFQVEAGALGPLKHEDWGTPPSAIDHCLSLVADEY 240
Db 182 MLRAAESYANMLGFPGKGSNEEGLEBEQIKAGALGLKVHEDWGSTPPAAINHALNVAQKY 241

Gy 241 DVQVCIHDTVNEAGYVDLTLANMNGRAHAHVHIEGAGCGHSPOVTWAGELNILPSSTT 300
Db 242 DVQAIIHTDLNEAGCVSDTWMKAIDGRTHIFTGAGGCHAPDIKKAAGENILPASFN 301

Gy 301 PTIPYNTVAEHLDMLTMCHHLDRKIREDLQFSQSRIIRPGSIAAEVLDHMGVIAMTSS 360
Db 302 PTIPTKNTADEHLDMLVMVCHLLDKKIKEDVAFADSRIIRPETIAAEDTLHDHGIPSITTSS 361

Gy 361 DSQAWGRAGEVIPRWQTADNNKKEFGKLPEP-DGHNDNNFRISKYISKYTIINPALTHGVS 419
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:

Db 362 DSQAMRGVGEVITRTWQTADCKNEFGALKKEBCG-ENDNFRIKRYISKYTINPAIAHGHS 420
 Qy 420 EYIGSVBECKIADLVVWNPAPFGVDPKIVKGMVVFSEMGDSNASVTPPQPVYTRMF 479
 Db 421 EYIGSVBECKIADLVVWNPAPFGVDPKIVKGMVVFSEMGDSNASVTPPQPVYTRMF 480
 Qy 480 HHGKAKPDTSTFVSKVAYENGVEKGLERQVLVVKNCNITKKDPKFNDAKTITVDP 539
 Db 481 SYGKAKNCALTFVSKYADCHKEBELGLERILLVVKNCNITKKDPKFNDAKTITVDP 540
 Qy 540 KTFEYFVDG-KLCTSKPTSQVPLAQRYTF 568
 Db 541 EYEVVRVNTKI-TSKPVEKVSGLQYCLF 569

RESULT 14
 URE1_BACSB
 ID URE1_BACSB STANDARD; PRT; 569 AA.
 AC Q07397, 1994, integrated into UniProtKB/Swiss-Prot.
 DT 01-OCT-1994, sequence version 1.
 DT 07-FEB-2006, entry version 51.
 GN Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase).
 GN Name-ureC;
 OS Bacillus sp. (strain TB-90).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=36824;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PARTIAL PROTEIN SEQUENCE.
 RX MEDLINE=94117379; PubMed=8288539; Masaki H., Uozumi T.;
 RA Maeda M., Hidaka M., Nakamura A., Masaki H., Uozumi T.;
 RT "Cloning, sequencing, and expression of thermophilic Bacillus sp.
 RT strain TB-90 urease gene complex in Escherichia coli.";
 RL J. Bacteriol. 176:432-442(1994).
 CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
 CC -1- SUBUNIT: Binds 2 nickel ions per subunit (By similarity).
 CC -1- COFACTOR: (Alpha, beta, gamma) (3) (By similarity).
 CC -1- PFM: Lys-219 is carbamylated. The carbamoyl group provides the
 CC ligands for the two nickel ions (By similarity).
 CC -1- SIMILARITY: Belongs to the urease family.

 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL: D14439; BAA03325.1; -; Genomic_DNA.
 DR PIR: C36950; C36950.
 DR HSP: P18314; 1EJX.
 DR InterPro: IPR006680; Amidohydro 1.
 DR InterPro: IPR011550; Amidohydro like.
 DR InterPro: IPR005848; Urease_alpha.
 DR InterPro: IPR011612; Urease_alpha_N.
 DR InterPro: IPR008295; Urease_alpha1one.
 DR Pfam: PF01979; Amidohydro_1; 1.
 DR Pfam: PF00449; Urease_alpha; 1.
 DR PRSF: PRSF001226; Urease_alpha; 1.
 DR PRINTS: PR01752; UREASE.
 DR ProDom: PD001248; Amidohydro like; 1.
 DR TIGRFAMs: TIGR01792; urease_alpha; 1.
 DR PROSITE: PS01120; UREASE_1; 1.
 DR PROSITE: PS00145; UREASE_2; 1.
 KW Direct protein sequencing; Hydrolase; Metal-binding; Nickel.
 FT CHAIN 1 569
 FT ACT_SITE 322 322
 FT METAL 136 136
 FT METAL 138 138
 FT METAL 219 219
 FT METAL 219 219
 FT METAL 248 248
 FT METAL 274 274
 FT METAL 362 362
 FT BINDING 221 221
 SQ SEQUENCE 569 AA; 61397 MW; 38AA83133A31A49D CRC64;

Query Match 68.2%; Score 2044; DB 1; Length 569;
 Best Local Similarity 66.4%; Pred. No. 8.1e-37;
 Matches 378; Conservative 76; Mismatches 108; Indels 7; Gaps 4;

Qy 3 MKKQEVNTVYPTGPKDKVGLDGLMAEVEHVDVTTTGEELKFGAGKTIKREGMQ---SNS 59
 Db 5 MRKQVADMFPTGVDAILRLADSELFIETIKDYTTTGDEVKFGGKVIKRGMGHPLATS 64
 Qy 60 PDENTLDLVTNAMIIDYTIYKADIKIKNGKIHGIGKAGNMQDQGVSPHVMVGVGTEA 119
 Db 65 -DE-CVDLVLTNAIIVDYTIYKADIKIKNGMIASIGKAGNPLLMQGV--MVIGAADEV 120
 Qy 120 LAGEGMIITAGGIDSHFLSPQFPALANGVTTMFGGTPGVDTGNATITPCKNLH 179
 Db 121 IAAEGMIVTAGGIDAHIFICPOQIETALASGVTTMIGGTPGATGNATTCTGPGNIIH 180
 Qy 180 RMLRAAEYSMVNGLKGNSSSKKOLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADE 239
 Db 181 RMLQAAEFPINLGLKGNCSDEAPLKEQIEAGAVGLKUHEDWGSTAAAIIDTCLKVADR 240
 Qy 240 YDVQCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVITMAGELNLPSS 299
 Db 241 YDVQVAILHTDLNEGGFVEDTLKAIDGRVHTVHTGAGGGHAPDIIKAAGFPNLPSS 300
 Qy 300 TPTPIYNTVAEHLDMLTCHHLDKRIREDLQFOSRIRPGSIAABDVLDHMGVIAMTS 359
 Db 301 NPTRPYNTLEHLDMLMVCCHLDANIPEDIAFADSRIRKETIAABDVLDHDLGVFSMIS 360
 Qy 360 SDSQAMGRAGEVTPRTWQTADCKNEFGALKKEBCGKNDNFRIKRYISKYTINPAIAHG 419
 Db 361 SDSQAMGRAGEVTPRTWQTADCKNEFGALKKEBCGKNDNFRIKRYISKYTINPAIAHG 420
 Qy 420 EYIGSVBECKIADLVVWNPAPFGVDPKIVKGMVVFSEMGDSNASVTPPQPVYTRMF 479
 Db 421 DVGSEVBECKIADLVVWNPAPFGVDPKIVKGMVVFSEMGDSNASVTPPQPVYTRMF 480
 Qy 480 HHGKAKPDTSTFVSKVAYENGVEKGLERQVLVVKNCNITKKDPKFNDAKTITVDP 539
 Db 481 AKGDAKYQTSITFVSKVAYENGVEKGLERQVLVVKNCNITKKDPKFNDAKTITVDP 540
 Qy 540 KTFEYFVDGKLCTSKPTSQVPLAQRYTF 568
 Db 541 QTYEVKVDQLVTCBPABIVPMAQRYFLF 569

RESULT 15
 Q9KG59_BACHD
 ID Q9KG59_BACHD PRELIMINARY; PRT; 571 AA.
 AC Q9KG59;
 DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2000, sequence version 1.
 DE Urease alpha subunit (EC 3.5.1.5).
 GN Name-ureC; OrderedLocusNames=BH0254;
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL: BA000004; BAB03973.1; -; Genomic_DNA.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:30:32 ; Search time 194 Seconds
(without alignments)
615.121 Million cell updates/sec

Title: US-11-056-825-2
Perfect score: 261
Sequence: 1 MAQVQLVQSGAEVKPGASV.....KVIKRGGLGLVYKDDDDK 261

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2589679 seqs, 457216429 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2589342

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_8.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*
10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	198	75.9	260	9 AEC92007	Aec92007 Single ch
2	51	19.5	110	8 ADW24449	Adw24449 HCMV-Pab-
3	48	18.4	93	8 ADU04216	Adu04216 Human imm
4	48	18.4	100	6 AAE35907	Aae35907 Human Al7
5	48	18.4	100	6 ABO27142	Abo27142 Human ger
6	48	18.4	100	7 ADF10189	Adf10189 Antibody
7	48	18.4	100	7 ADF10087	Adf10087 VEGF anti
8	48	18.4	100	7 ADF09982	Adf09982 Antibody
9	48	18.4	100	7 ADJ80262	Adj80262 V kappa ge
10	48	18.4	100	8 ADO07331	Ado07331 Human ant
11	48	18.4	100	9 ADY75436	Ady75436 Human ger
12	48	18.4	100	9 AEB13589	Aeb13589 Human var
13	48	18.4	100	9 AEB13642	Aeb13642 Human var
14	48	18.4	100	9 AED04308	Aed04308 Human ant
15	48	18.4	100	9 AED06937	Aed06937 Human ant
16	48	18.4	100	10 AEA47598	Aea47598 Human CDR
17	48	18.4	100	10 AEF18501	Aef18501 HABA prot
18	48	18.4	111	4 AAE06948	Aae06948 Human HF-
19	48	18.4	111	4 AAE06996	Aae06996 Human kap
20	48	18.4	111	4 AAU09920	Aau09920 Human ant
21	48	18.4	111	7 ABR61476	Adr61476 Human ant
22	48	18.4	111	8 ADQ89233	Adq89233 Human imm
23	48	18.4	111	8 ADQ89281	Adq89281 Human imm

24	48	18.4	111	9 ADU26553	Adu26553 Human ant
25	48	18.4	111	9 AEB09554	Aeb09554 Human lig
26	48	18.4	111	9 AEB09506	Aeb09506 Human HF-
27	48	18.4	111	9 AEC92164	Aec92164 Human HF-
28	48	18.4	112	2 AAW79227	Aaw79227 Light cha
29	48	18.4	112	2 AAW53585	Aaw53585 Light cha
30	48	18.4	112	4 AAE07003	Aae07003 Human kap
31	48	18.4	112	4 AAE06993	Aae06993 Human kap
32	48	18.4	112	4 AAE06999	Aae06999 Human kap
33	48	18.4	112	5 ADF98232	Adf98232 Human HF-
34	48	18.4	112	8 ADQ89284	Adq89284 Human imm
35	48	18.4	112	8 ADQ89278	Adq89278 Human imm
36	48	18.4	112	8 ADQ89288	Adq89288 Human imm
37	48	18.4	112	9 ADX01284	Adx01284 Human Hu5
38	48	18.4	112	9 AEB09557	Aeb09557 Human lig
39	48	18.4	112	9 AEB09551	Aeb09551 Human lig
40	48	18.4	112	9 AEB09561	Aeb09561 Human lig
41	48	18.4	113	8 ADO07283	Ado07283 Human pro
42	48	18.4	113	9 ADZ42043	Adz42043 Ig L chai
43	48	18.4	113	9 ADZ42042	Adz42042 Ig L chai
44	48	18.4	113	9 ADZ42041	Adz42041 Ig L chai
45	48	18.4	121	10 AEG01367	Aeg01367 Kallikrei

ALIGNMENTS

RESULT 1
AEC92007
ID AEC92007 standard; protein; 260 AA.
XX
AC AEC92007;
XX
DT 01-DEC-2005 (first entry)
XX
DE Single chain variable fragment antibody BC-12 protein SEQ ID NO 2.
XX

cytostatic; immunostimulant; pharmaceutical; protein interaction;
therapeutic; immune modulation; solid tumor; leukemia; colorectal
tumor;
breast tumor; uterine cervix tumor; uterine fibroids; ovary tumor;
gynecology and obstetrics; polycystic ovary syndrome; antiinfertility;
gynecological; polyp; growth disorder; prostate tumor; prostate disease;
androgyny; genitourinary disease; pituitary tumor; endocrine disease;
KW adenocarcinoma; melanoma; bone tumor; musculoskeletal disease;
KW multiple myeloma; hematological disease; immune disorder;
KW central nervous system tumor; neurological disease; glioma;
KW astroblastoma; neoplasm; single chain antibody; BC-12.

OS Homo sapiens.
OS Synthetic.

Key Location/Qualifiers
FT Misc-difference 1..2 "Encoded by ATG"
FT Misc-difference 62
/note= "Encoded by GCA"

PN WO2005091805-A2.

XX 06-OCT-2005.

PD 14-FEB-2005; 2005WO-US004612.

XX 13-FEB-2004; 2004US-0544807P.

PR 10-NOV-2004; 2004US-0626726P.

PR 11-FEB-2005; 2005US-00056825.

XX (SCRI) SCRIPPS RES INST.

XX Felding-Habermann B, Janda KD, Saven A;

XX WPI; 2005-664872/68.

DR N-PSDB; AEC92006.

XX New scFv Bc-12 or Bc-15 antibody comprises ligand mimetic, which
PT specifically binds to activated alphavbeta3 integrin receptor, useful for
PT treating, e.g. neoplastic disease, solid tumor, leukemia, or benign or
PT malignant breast cancer.
XX
PS Claim 2; SEQ ID NO 2; 109pp; English.
XX
CC The invention describes an antibody comprising a ligand mimetic, which
CC specifically binds to an activated alphavbeta 3 integrin receptor
CC differentially produced on a cell in a metastatic state compared to a
CC similar, non-metastatic cell, and does not bind to a non-activated
CC alphavbeta 3 integrin receptor. Also described are: a pharmaceutical
CC composition comprising the antibody; treating a disease state in a mammal
CC; a cell line comprising a tumor cell variant with a metastatic homing
CC propensity to a target tissue; producing an antibody phage population
CC having affinity for a tumor cell target; detecting tumor cells in a
CC mammal by treatment with a cancer therapeutic; inducing or enhancing an
CC immune response to an antigen in a mammal; detecting an activated cell
CC surface receptor on a metastatic tumor cell surface in a mammalian tissue
CC sample; interfering with cells liable to undergo metastasis associated
CC with a disease state; identifying cells liable to undergo metastasis
CC associated with a disease state; an isolated Bc-12 or Bc-15
CC polynucleotide comprising a nucleotide sequence that has at least 90%
CC percent identity to a fully defined 802 or 752 bp sequence (SEQ ID NO. 1
CC or 3) given in the specification; an isolated polypeptide comprising a
CC nucleotide sequence that has at least 90% sequence identity to SEQ ID NO.
CC 1 or 3, or shares a biological function with Bc-12 or Bc-15; a vector
CC comprising the polynucleotide; an expression vector comprising the
CC polynucleotide in which the nucleotide sequence of the polynucleotide is
CC operatively linked with a regulatory sequence that controls expression of
CC the polynucleotide in a host cell; a host cell comprising the
CC polynucleotide or progeny of the cell; and determining anti-metastatic
CC activity of a test compound in a mammal. The antibody is useful for
CC treating neoplastic disease, solid tumor, hematological malignancy,
CC leukemia, colorectal cancer, benign or malignant breast cancer, uterine
CC cancer, uterine leiomyomas, ovarian cancer, endometrial cancer,
CC polycystic ovary syndrome, endometrial polyps, prostate cancer, prostatic
CC hypertrophy, pituitary cancer, adenomyosis, adenocarcinomas, meningioma,
CC melanoma, bone cancer, multiple myeloma, CNS cancer, glioma, or
CC astroblastoma. This is the amino acid sequence of single chain variable
CC fragment antibody Bc-12.
XX
SQ Sequence 260 AA;

Query Match 75.9%; Score 198; DB 9; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.8e-156; Indels 0; Gaps 0;
Matches 198; Conservative 0; Mismatches 0;
QY 62 QKFGQRTMTTDTSTAYMELSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGGG 121
DB 63 QKFGQRTMTTDTSTAYMELSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGGG 122
QY 122 GSGGGGGGGGSEIVLTQSPLSPLVTLLQCPASISCRSSQNLVYSDGNTYLSWFOQRPQS 181
DB 123 GSGGGGGGGGSEIVLTQSPLSPLVTLLQCPASISCRSSQNLVYSDGNTYLSWFOQRPQS 182
QY 182 PRLIYKVSNRDGVDPFSGSGSGTDTLKISRVEADIGVYCMQGTTHWPPRTFGQGT 241
DB 183 PRLIYKVSNRDGVDPFSGSGSGTDTLKISRVEADIGVYCMQGTTHWPPRTFGQGT 242
QY 242 KVEIKRGLGLVYDKDD 259
DB 243 KVEIKRGLGLVYDKDD 260

RESULT 2
ADW24449
ID ADW24449 standard; protein; 110 AA.
XX
AC ADW24449;
XX
DT 10-MAR-2005 (first entry)

XX HCMV-Fab-6 amino acid sequence #1.
DE Humanized antibody; antibody engineering; fab; cytomegalovirus infection.
XX
XX Human herpesvirus 5.
PN CN1445243-A.
XX
PD 01-OCT-2003.
XX
PF 13-DEC-2002; 2002CN-00155426.
XX
PR 13-DEC-2002; 2002CN-00155426.
XX
PA (VIRU-) INST VIRUS PREVENTION & CONTROL CHINA DI.
XX
PI Liang M, Gu S, Duan T;
XX
XX WPI; 2004-091790/10.
DR N-PSDB; ADW24448.
XX
PT Fab antibody of human resourced neutrality genetic engineering for anti
XX human cytomegalo virus.
PS Disclosure; Fig 3b; 15pp; Chinese.
XX
CC The invention relates to a humanized neutral genetically engineering Fab
CC antibody (HCMV-Fab-2 and HCMV-Fab-6) of human cytomegalovirus. The Fab
CC fragment, gene product and application are also disclosed. The
CC recombinant antibody is determined by the specific gene sequence in the
CC hypervariable region of the light-chain and heavy-chain variable regions.
CC The antibody can be used for preventing and treating the diseases
CC associated with HCMV infection. The current sequence represents a HCMV-
CC Fab-6 amino acid sequence.
XX
SQ Sequence 110 AA;
Query Match 19.5%; Score 51; DB 8; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.6e-34;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 YLSWFOQRPQSGPRLLIYKVSNRDGVDPFSGSGSGTDTLKISRVEAD 220
DB 35 YLSWFOQRPQSGPRLLIYKVSNRDGVDPFSGSGSGTDTLKISRVEAD 85
RESULT 3
ADU04216
ID ADU04216 standard; protein; 93 AA.
XX
AC ADU04216;
XX
DT 13-JAN-2005 (first entry)
XX
DE Human immunoglobulin kappa light chain protein 5.
XX
KW viral infection; virucide; autoimmune disease;
KW lymphoproliferative disorder; vaccine; gene therapy; immunoglobulin;
KW antibody.
XX
OS Homo sapiens.
XX
PN WO2004090544-A2.
XX
PD 21-OCT-2004.
XX
PF 13-APR-2004; 2004WO-CA000544.
XX
PR 09-APR-2003; 2003US-0461137P.
PR 30-SEP-2003; 2003US-0506779P.
XX
PA (CABL-) CANADIAN BLOOD SERVICES.

XX HU Y, Brown B;
XX WPI; 2004-766498/75.
XX
XX Characterizing a viral infection in a host, for developing treatment for
XX severe acute respiratory syndrome-coronavirus (SARS-CoV), by determining
XX homology profile of a viral-based sequence element with an endogenous
XX host element.
XX
XX Example 1; Fig 24; 166pp; English.
XX
XX The invention relates to a novel method for characterising a viral
XX infection in a host. The method comprises identifying at least one viral-
XX based sequence element in a biological sample obtained from the host,
XX determining a homology profile of the viral-based sequence element with
XX at least one endogenous host element and characterising the viral
XX infection based on the homology profile, where the homology profile is
XX indicative of a viral behaviour of the viral infection in the host. The
XX method of the invention demonstrates virucide applications and may be
XX useful for preparing a medicament for detecting and/or treating a viral
XX infection or related condition, such as an autoimmune disease e.g. type
XX II cryoglobulinemia, or lymphoproliferative disorder. The viral
XX infection may be due to HCV (Hepatitis C virus), HIV or a member of a
XX Retroviridae, Flaviviridae, Herpesviridae, Papillomaviridae or
XX Coronaviridae virus family. Treatment of the infection may utilise
XX vaccine or gene therapy. The target compound of the invention may be
XX useful for detecting the presence of a virus in a biological sample or
XX for manufacturing a medicament for treating SARS coronavirus (SARS-CoV)
XX and/or Human T-lymphotropic virus 1 (HTLV-I) infection. The methods are
XX further useful for developing treatment regimes to target genotype-
XX specific viral variants. The current sequence is that of a human
XX immunoglobulin light chain kappa (IgVlk) protein of the invention.
XX
XX Sequence 93 AA;
XX
XX Query Match 18.4%; Score 48; DB 8; Length 93;
XX Best Local Similarity 100.0%; Pred. No. 7.1e-32;
XX Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 173 WFOQRPQSGSPRLIYKVSNRDSGVPDRFSGSGGTDTFTLKISRVEAED 220
XX |||||
XX Db 40 WFOQRPQSGSPRLIYKVSNRDSGVPDRFSGSGGTDTFTLKISRVEAED 87
XX
XX
XX RESULT 4
XX AAE35907
XX ID AAE35907 standard; protein; 100 AA.
XX AC AAE35907;
XX DT 17-JUN-2003 (first entry)
XX DE Human A17 antibody light chain germline protein.
XX KW Human; cytotoxic T lymphocyte antigen-4; therapy; CTLA-4 antibody;
XX cancer.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 24..39
XX /note= "Complementarity determining region (CDR) 1"
XX Region 55..61
XX /note= "Complementarity determining region (CDR) 2"
XX Region 94..100
XX /note= "Complementarity determining region (CDR) 3"
XX
XX EP1262193-A1.
XX
XX 04-DEC-2002.
XX
XX 23-MAY-2002; 2002EP-00253652.

XX 23-MAY-2001; 2001US-0293042P.
XX (PFIZ) PFIZER PROD INC.
XX Hanson DC, Mueller EB;
XX WPI; 2003-131215/13.
XX
XX Use of human anti-cytotoxic T lymphocyte antigen-4 antibody in the
XX preparation of medicament for the treatment of cancer.
XX
XX Disclosure; Fig 7; 76pp; English.
XX
XX The invention relates to the use of human anti-cytotoxic T lymphocyte
XX antigen-4 (anti-CTLA-4) antibody in the preparation of a medicament for
XX the treatment of cancer such as bone cancer, pancreatic cancer, skin
XX cancer, cutaneous or intraocular malignant melanoma, uterine and ovarian
XX cancer, cancer of the anal region, stomach cancer, breast cancer,
XX testicular cancer, uterine cancer and carcinoma of the fallopian tubes.
XX The present sequence is human A10/A26 antibody light chain germline
XX protein. This sequence is used in the invention
XX
XX Sequence 100 AA;
XX
XX Query Match 18.4%; Score 48; DB 6; Length 100;
XX Best Local Similarity 100.0%; Pred. No. 7.5e-32;
XX Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 173 WFOQRPQSGSPRLIYKVSNRDSGVPDRFSGSGGTDTFTLKISRVEAED 220
XX |||||
XX Db 40 WFOQRPQSGSPRLIYKVSNRDSGVPDRFSGSGGTDTFTLKISRVEAED 87
XX
XX
XX RESULT 5
XX ABO27142
XX ID ABO27142 standard; protein; 100 AA.
XX AC ABO27142;
XX DT 10-SEP-2003 (first entry)
XX DE Human germline light chain variable region gene segment #22.
XX KW Human; light chain variable region; VK; humanised antibody;
XX chimeric antibody; complementarity determining region; CDR;
XX canonical CDR structure type.
XX OS Homo sapiens.
XX US2003039649-A1.
XX 27-FEB-2003.
XX
XX 12-JUL-2002; 2002US-00194975.
XX
XX 12-JUL-2001; 2001US-0305111P.
XX (FOOT/) FOOTE J.
XX
XX Foote J;
XX
XX WPI; 2003-492151/46.
XX
XX Making humanized antibody for converting antibody, by making chimeric
XX antibodies containing complementarity determining region from non-human
XX antibody and appropriate framework sequences of human antibodies.
XX
XX Example 1; Fig 2; 31pp; English.
XX
XX The invention describes a method of making a humanised antibody,
XX comprising making chimeric antibodies containing a complementarity
XX determining region (CDR) from a non-human antibody and appropriate

framework sequences (I) of human antibodies. (I) is selected by using canonical CDR structure types of non-human antibody in comparison to germline canonical CDR structure types of human antibodies as the basis for selection, for humanisation. The method is useful for making a humanised antibody or a converted antibody. The method is applicable for converting a subject antibody sequence of any subject species to a less immunogenic form suitable for use in an object species. The method is reliable for identifying suitable human framework sequences to support non-human CDR regions and to provide humanised antibodies that retain high antigen binding with low immunogenicity in humans, without the need for direct comparison of framework sequences, without the need for determining critically important amino acid residues in the framework, and without the need for multiple iteration and construction to obtain humanised antibodies with suitable therapeutic properties. The antibody has high affinity and low immunogenicity without need for comparing framework sequences between non-human and human antibodies. This sequence represents a human light chain variable region gene segment used in the creation of humanised antibodies

XX Sequence 100 AA;

Query Match 18.4%; Score 48; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGTDTFLKISRVEAED 220
|||||
Db 40 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGTDTFLKISRVEAED 87

RESULT 6

ADFI10189
ID ADFI10189 standard; protein; 100 AA.

XX ADFI10189;

DT 12-FEB-2004 (first entry)

XX Antibody light chain variable region VLK_2-30.

XX Antibody; stability; solubility; antigen binding affinity;
XX variable region; human.

XX Homo sapiens.

XX WO2003074679-A2.

XX 12-SEP-2003.

XX 03-MAR-2003; 2003WO-US006598.

XX 01-MAR-2002; 2002US-0360843P.

XX 29-MAY-2002; 2002US-0384197P.

XX (XENC-) XENCOR.

XX Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

XX WPI; 2003-722066/68.

XX Computer optimization of physicochemical properties of antibodies
XX comprises analyzing the interactions of amino acids at variable
XX positions.

XX Example 16; Fig 40b; 135pp; English.

XX The present invention relates to a method for optimizing at least one
XX physico-chemical property of an antibody by a computational screening
XX method. The method comprises: receiving a template antibody structure;
XX selecting at least one variable position belonging to the antibody
XX structure; selecting at least one amino acid to be considered at the
XX variable position(s); analyzing the interaction of each selected amino
XX acid at each variable position with at least part of the remainder of the

CC antibody, including the selected amino acids at other variable positions;
CC and identifying a set of at least one antibody sequence with at least one
CC optimized physico-chemical property. The method is useful for optimizing
CC the physico-chemical properties of an antibody, especially the stability,
CC solubility, or antigen binding affinity. The optimized antibody may be
CC useful for treating a patient. The present sequence is an antibody
CC variable region sequence used to illustrate the invention.

XX Sequence 100 AA;

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGTDTFLKISRVEAED 220
|||||
Db 40 WFOQRPGQSPRLIYKVSNRDGVDPFRSGSGTDTFLKISRVEAED 87

RESULT 7

ADFI10087
ID ADFI10087 standard; protein; 100 AA.

XX ADFI10087;

DT 12-FEB-2004 (first entry)

XX VEGF antibody light chain variable region VLK_2-30.

XX Antibody; stability; solubility; antigen binding affinity;
XX variable region; human; VEGF.

XX Homo sapiens.

XX WO2003074679-A2.

XX 12-SEP-2003.

XX 03-MAR-2003; 2003WO-US006598.

XX 01-MAR-2002; 2002US-0360843P.

XX 29-MAY-2002; 2002US-0384197P.

XX (XENC-) XENCOR.

XX Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

XX WPI; 2003-722066/68.

XX Computer optimization of physicochemical properties of antibodies
XX comprises analyzing the interactions of amino acids at variable
XX positions.

XX Example 6; Fig 16b; 135pp; English.

XX The present invention relates to a method for optimizing at least one
XX physico-chemical property of an antibody by a computational screening
XX method. The method comprises: receiving a template antibody structure;
XX selecting at least one variable position belonging to the antibody
XX structure; selecting at least one amino acid to be considered at the
XX variable position(s); analyzing the interaction of each selected amino
XX acid at each variable position with at least part of the remainder of the
XX antibody, including the selected amino acids at other variable positions;
XX and identifying a set of at least one antibody sequence with at least one
XX optimized physico-chemical property. The method is useful for optimizing
XX the physico-chemical properties of an antibody, especially the stability,
XX solubility, or antigen binding affinity. The optimized antibody may be
XX useful for treating a patient. The present sequence is an antibody
XX variable region sequence used to illustrate the invention.

XX Sequence 100 AA;

Query Match 18.4%; Score 48; DB 7; Length 100;

Best Local Similarity 100.0%; Pred. No. 7.5e-32; Mismatches 0; Indels 0; Gaps 0;
Matches 48; Conservative 0;

QY 173 WFOQRPQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 220
DB 40 WFOQRPQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 87

RESULT 8
ADFO9982
ID ADF09982 standard; protein; 100 AA.
XX
AC ADF09982;
XX
DT 12-FEB-2004 (first entry)
XX
DE Antibody light chain variable region VLK_2-30.
XX
DE Antibody; stability; solubility; antigen binding affinity;
KW variable region; human.
XX
OS Homo sapiens.
XX
PN WO2003074679-A2.
XX
PD 12-SEP-2003.
XX
PF 03-MAR-2003; 2003WO-US006598.
XX
PR 01-MAR-2002; 2002US-0360843P.
PR 29-MAY-2002; 2002US-0384197P.
XX
PA (XENC-) XENCOR.
XX
PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;
XX
DR WPI; 2003-722066/69.
XX
PT Computer optimization of physicochemical properties of antibodies
PT comprises analyzing the interactions of amino acids at variable
PT positions.
XX
PS Disclosure; Fig 2b; 135pp; English.
XX
XX The present invention relates to a method for optimizing at least one
XX physico-chemical property of an antibody by a computational screening
XX method. The method comprises: receiving a template antibody structure;
XX selecting at least one variable position belonging to the antibody
XX structure; selecting at least one amino acid to be considered at the
XX variable position(s); analyzing the interaction of each selected amino
XX acid at each variable position with at least part of the remainder of the
XX antibody, including the selected amino acids at other variable positions;
XX and identifying a set of at least one antibody sequence with at least one
XX optimized physico-chemical property. The method is useful for optimizing
XX the physico-chemical properties of an antibody, especially the stability,
XX solubility, or antigen binding affinity. The optimized antibody may be
XX useful for treating a patient. The present sequence is an antibody
XX variable region sequence used to illustrate the invention.

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32; Mismatches 0; Indels 0; Gaps 0;
Matches 48; Conservative 0;

QY 173 WFOQRPQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 220
DB 40 WFOQRPQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 87

RESULT 9
ADJ80262
ID ADJ80262 standard; protein; 100 AA.

XX ADJ80262;
XX
DT 06-MAY-2004 (first entry)
XX
DE Vkappa gene locus antibody amino acid sequence #22.
XX
KW hybrid antibody; antibody; framework region; homology; immunogenicity.
XX
OS Homo sapiens.
XX
PN WO2003048321-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-US038450.
XX
PR 03-DEC-2001; 2001US-0336591P.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Rother R, Wu D;
XX
DR WPI; 2003-513753/48.
XX
PT Producing a hybrid antibody or hybrid antibody fragment by operatively
PT linking the selected framework sequences to one or more complementarity
PT determining regions of the initial antibody.
XX
PS Disclosure; SEQ ID NO 22; 77pp; English.
XX
XX The invention relates to a method of producing a hybrid antibody or
XX hybrid antibody fragment by: (i) providing an initial antibody having
XX specificity for a target; (ii) determining the sequence of a variable
XX region of the initial antibody; (iii) selecting a first component of the
XX variable region consisting of FR1, FR2, FR3 and FR4; (iv) comparing the
XX sequence of the first component to sequences contained in a reference
XX database of antibody sequences or antibody fragment sequences from a
XX target species; (v) selecting a sequence from an antibody in the database
XX which demonstrates a high degree of homology to the first component; (vi)
XX selecting a second component of the variable region which is different
XX than the first component, the second component selected from the group
XX consisting of FR1, FR2, FR3 and FR4; (vii) comparing the sequence of the
XX second component to sequences contained in a reference database of
XX antibody sequences or antibody fragment sequences from the target species
XX; (viii) selecting a sequence from the database which demonstrates a high
XX degree of homology to the second component and which is from a different
XX antibody than the selected antibody; and (ix) operatively linking the
XX selected framework sequences to one or more complementarity determining
XX regions (CDRs) of the initial antibody to produce a hybrid antibody or
XX hybrid antibody fragment. The method is useful for producing a hybrid
XX antibody or hybrid antibody fragment (claimed). The antibody and
XX fragments are useful for therapeutic and diagnostic purposes. The method
XX uses entire framework regions from a single antibody variable heavy or
XX variable light chain to receive the CDRs. This produces antibodies that
XX are highly homologous and exhibit reduced immunogenicity while
XX maintaining an optimum binding profile. This sequence represents the
XX amino acid sequence of an antibody from the Vkappa gene locus.

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32; Mismatches 0; Indels 0; Gaps 0;
Matches 48; Conservative 0;

QY 173 WFOQRPQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 220
DB 40 WFOQRPQSPRLIYKVSNRDSGVDPFRFSGSGGTDFTLKISRVEAED 87

RESULT 10
ADO07331
ID ADO07331 standard; protein; 100 AA.

XX ADO07331;
 XX 15-JUL-2004 (first entry)
 XX Human antibody A17 light chain variable region.
 XX Catalytic antibody; human; antibody; gene therapy; protease.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX Region 1. .23
 XX /label= FR1
 XX /note= "Framework region 1"
 XX Active-site 1
 XX Region 24. .39
 XX /label= CDR1
 XX /note= "Complementarity determining region 1"
 XX Active-site 25
 XX /note= "Possible Ser component of catalytic triad"
 XX Active-site 26
 XX /note= "Possible Ser component of catalytic triad"
 XX Active-site 28
 XX /note= "Possible Ser component of catalytic triad"
 XX Active-site 32
 XX /note= "Possible Ser component of catalytic triad"
 XX Active-site 33
 XX /note= "Possible Asp component of catalytic triad"
 XX Region 40. .54
 XX /label= FR2
 XX /note= "Framework region 2"
 XX Region 55. .61
 XX /label= CDR2
 XX /note= "Complementarity determining region 2"
 XX Active-site 57
 XX /note= "Possible Ser component of catalytic triad"
 XX Active-site 60
 XX /note= "Possible Asp component of catalytic triad"
 XX Active-site 61
 XX /note= "Possible Ser component of catalytic triad"
 XX Region 62. .93
 XX /label= FR3
 XX /note= "Framework region 3"
 XX Region 94. .100
 XX /label= CDR3
 XX /note= "Complementarity determining region 3"
 XX Active-site 98
 XX /note= "Possible His component of catalytic triad"
 XX WO2004033658-A2.
 XX 22-APR-2004.
 XX 09-OCT-2003; 2003WO-US032214.
 XX 10-OCT-2002; 2002US-0417979P.
 XX (INTE-) INTEGRIGEN INC.
 XX Smider V, Larrick JW;
 XX WPI; 2004-340921/31.
 XX Novel recombinant catalytic polypeptide useful for cleaving target
 XX proteins or for treating or preventing cancers, comprises a human
 XX antibody light chain operably joined to a heterologous antibody heavy
 XX chain.
 XX Disclosure; Fig 3; 92pp; English.
 XX The present sequence is the light chain variable region of human antibody

CC A17. This is one of a repertoire of human kappa light chain sequences
 CC ADO07310-ADO07349 screened for putative catalytic triads. Several genes
 CC encoding such light chains ADO07282-ADO07309 were cloned for use in
 CC recombinant catalytic polypeptides of the invention. These comprise a
 CC human antibody light chain operably joined to a heterologous antibody
 CC heavy chain. The light chain has a serine protease dyad and endopeptidase
 CC activity, and the heavy chain has a predetermined specificity for a
 CC target protein. By joining 2 heterologous human antibody chains, one of
 CC which supplies the catalytic activity to hydrolyse polypeptides and the
 CC other the binding specificity for a target protein, the invention
 CC provides for the construction of a repertoire of proteases with
 CC customised protein substrate specificities of potentially unlimited
 CC number and thus makes possible the effective treatment and/or prevention
 CC of any medical condition attributable to the presence or overexpression
 CC of an identified protein. The invention also provides nucleic acids
 CC encoding the catalytic antibodies (which can be used for gene therapy),
 CC host cells, transgenic non-human animals, and methods of cleaving a
 CC target protein (in vitro or in vivo) using a recombinant catalytic
 CC polypeptide. It also provides a library of recombinant catalytic
 CC polypeptides with altered enzymatic activity, and a method of altering
 CC the enzymatic activity of the recombinant catalytic polypeptides by
 CC mutating at least one complementarity determining region of the heavy
 CC chain. A serine protease triad was identified in the present sequence,
 CC which can be used as the light chain sequence in catalytic polypeptides
 CC of the invention ADO07283.
 XX SQ Sequence 100 AA;
 Query Match 18.4%; Score 48; DB 8; Length 100;
 Best Local Similarity 100.0%; Pred. No. 7.5e-32; Mismatches 0; Gaps 0;
 Matches 48; Conservative 0; Indels 0;
 QY 173 WFQRPQSGSPRLIYKVSNRDGVDPFRFSGSGGTDFTLKISRVEAED 220
 DB 40 WFQRPQSGSPRLIYKVSNRDGVDPFRFSGSGGTDFTLKISRVEAED 87
 RESULT 11
 ADY75436
 ID ADY75436 standard; peptide; 100 AA.
 XX AC ADY75436;
 XX 02-JUN-2005 (first entry)
 XX Human germline light chain kappa V minigene VKII peptide #3.
 XX Antibody engineering; antibody; antibody production; gene library;
 XX DNA recombination; gene amplification; primer extension;
 XX light chain variable region.
 XX Homo sapiens.
 XX WO2005023993-A2.
 XX 17-MAR-2005.
 XX 09-SEP-2004; 2004WO-US029617.
 XX 09-SEP-2003; 2003US-0501073P.
 XX (INTE-) INTEGRIGEN INC.
 XX Sharma V, Leonard L, Smider V;
 XX WPI; 2005-223364/23.
 XX Producing polynucleotide encoding human germline antibody V-region for
 XX generating full-length antibody germline V-region genes, by obtaining V
 XX or J minigene and joining V minigene with J minigene, or joining J
 XX minigene with V minigene.
 XX Disclosure; Fig 12; 52pp; English.

XX The present invention relates to producing germline antibody genes by a
CC completely in vitro approach that mimics the natural process of V(D)J
CC recombination. The antibody genes are completely human and native in
CC their sequence, and libraries of such antibody genes can be constructed
CC which represent an unselected population representing the entire antibody
CC repertoire. The method uses gene amplification to produce a V minigene,
CC and a hybrid primer capable of hybridizing to a V minigene and either a D
CC or V minigene. The hybrid primer facilitates recombination of a V
CC minigene to a D or J minigene to produce a full length V-region gene.
CC Also disclosed is a library comprising member polynucleotides encoding a
CC exogenously rearranged human germline antibody V-regions. In producing a
CC polynucleotide encoding a human germline antibody V-region, a D minigene
CC is further joined to the 3' end of the V minigene and the 5' end of the J
CC minigene. The V minigene or the J minigene in is obtained by chemical
CC synthesis or by amplification from a germline DNA library. Joining the V
CC minigene with at least one J minigene is performed by primer extension
CC using at least two or three oligonucleotide primers. The V minigene is
CC derived from human immunoglobulin kappa locus, human immunoglobulin
CC lambda locus, or human immunoglobulin heavy chain locus. The V-region
CC also comprises a serine protease triad. The human germline antibodies can
CC be used as precursors to more high affinity antibodies, and are useful in
CC the generation of efficiently pairing libraries of heavy and light
CC chains. The present sequence is a human germline light chain kappa region
CC peptide.
XX
SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQSPRLIYKVNDSGVDPDRFSGSGGTDTFLKISRVEAD 220
DB 40 WFOQRPQSPRLIYKVNDSGVDPDRFSGSGGTDTFLKISRVEAD 87

RESULT 12
AEB13589
ID AEB13589 standard; protein; 100 AA.
AC AEB13589;
DT 25-AUG-2005 (first entry)
DE Human variable heavy chain region protein, VH_3-43.
XX protein engineering; immunogenicity; germ cell;
KW heavy chain variable region; antibody.
XX Homo sapiens.
OS
PN WO2005056759-A2.
XX
PD 23-JUN-2005.
XX
PF 03-DEC-2004; 2004WO-US040694.
XX
PR 04-DEC-2003; 2003US-0527167P.
PR 21-JUN-2004; 2004US-0581613P.
PR 13-AUG-2004; 2004US-0601665P.
PR 16-OCT-2004; 2004US-0619483P.
XX
PA (XENC-) XENCOR INC.
XX
PI Lazar GA, Desjarlais JR, Hammond PW;
XX WPI; 2005-458579/46.
XX
PT Generating variant protein for host, by comparing parent protein sequence
PT with natural protein sequences from host, analyzing and substituting
PT amino acids of parent sequences with corresponding amino acid string of
PT natural sequence.

XX Disclosure; Fig 1a; 137pp; English.
XX
CC The invention relates to a novel method for generating a variant protein
CC for a host. The method involves comparing a parent protein sequence with
CC natural protein sequences from a host, analyzing the amino acid strings
CC of the parent sequence with corresponding amino acid strings of each of
CC the natural protein sequences, and substituting amino acids of parent
CC protein sequences with corresponding amino acid strings of a natural
CC protein sequence on an amino acid string. The method is useful for
CC generating a variant protein, e.g. a variant antibody for a host as
CC compared to a parent protein. The method enables the generation of
CC variant proteins having an increased host string content and reduced
CC immunogenicity. This sequence represents a variable heavy chain of a
CC human antibody protein used in a human germ line comparison of the
CC invention.
XX
SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQRPQSPRLIYKVNDSGVDPDRFSGSGGTDTFLKISRVEAD 220
DB 40 WFOQRPQSPRLIYKVNDSGVDPDRFSGSGGTDTFLKISRVEAD 87

RESULT 13
AEB13642
ID AEB13642 standard; protein; 100 AA.
XX
AC AEB13642;
DT 25-AUG-2005 (first entry)
XX
DE Human variable light kappa chain region protein, Vlk_2-30.
XX
KW protein engineering; immunogenicity; germ cell;
KW light chain variable region; antibody.
XX
OS Homo sapiens.
XX
PN WO2005056759-A2.
XX
PD 23-JUN-2005.
XX
PF 03-DEC-2004; 2004WO-US040694.
XX
PR 04-DEC-2003; 2003US-0527167P.
PR 21-JUN-2004; 2004US-0581613P.
PR 13-AUG-2004; 2004US-0601665P.
PR 16-OCT-2004; 2004US-0619483P.
XX
PA (XENC-) XENCOR INC.
XX
PI Lazar GA, Desjarlais JR, Hammond PW;
XX WPI; 2005-458579/46.
XX
PT Generating variant protein for host, by comparing parent protein sequence
PT with natural protein sequences from host, analyzing and substituting
PT amino acids of parent sequences with corresponding amino acid string of
PT natural sequence.

XX Disclosure; Fig 1a; 137pp; English.
XX
CC The invention relates to a novel method for generating a variant protein
CC for a host. The method involves comparing a parent protein sequence with
CC natural protein sequences from a host, analyzing the amino acid strings
CC of the parent sequence with corresponding amino acid strings of each of
CC the natural protein sequences, and substituting amino acids of parent
CC protein sequences with corresponding amino acid strings of a natural
CC protein sequence.

CC protein sequence on an amino acid string. The method is useful for
CC generating a variant protein, e.g. a variant antibody for a host as
CC compared to a parent protein. The method enables the generation of
CC variant proteins having an increased host string content and reduced
CC immunogenicity. This sequence represents a variable light chain of a
CC human antibody protein used in a human germ line comparison of the
CC invention.

XX
SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQRPQGSPPRLIYKVSNRDSGVPDRFSGSGGTFTLKISRVEAED 220
|||||
DB 40 WFQRPQGSPPRLIYKVSNRDSGVPDRFSGSGGTFTLKISRVEAED 87
|||||

RESULT 14
ID AED04308
XX AED04308 standard; protein; 100 AA.
AC AED04308;
XX
DT 01-DEC-2005 (first entry)
XX
DE Human anti-CTLA-4 antibody, kappa chain protein SEQ ID NO:36.
XX
KW cancer; cytostatic; neoplasm; antibody; antibody therapy; igg; CTLA-4;
KW cytotoxic T lymphocyte antigen-4.
XX
OS Homo sapiens.
XX
XX WO2005092380-A2.
PN
XX
PD 06-OCT-2005.
XX
XX 14-MAR-2005; 2005WO-IB000671.
XX
XX 26-MAR-2004; 2004US-0556801P.
XX
XX (PFIZ) PFIZER PROD INC.
XX
XX Gomez-Navarro J, Hanson DC, Mueller EE, Noe DA;
XX
XX WPI; 2005-684091/70.
XX
PT Treating cancer in mammal who has undergone stem cell transplantation,
PT comprises administering human anti-cytotoxic T lymphocyte antigen-4
PT antibody to the mammal.
XX
XX Disclosure; SEQ ID NO 36; 135pp; English.

XX The invention relates to a method of treating (M1) cancer in a mammal
CC comprising administration of a human anti-CTLA-4 (cytotoxic T lymphocyte
CC antigen-4) antibody. Also included are: treating cancer in a mammal by
CC administering more than 10 mg/ml of a human anti-CTLA-4 antibody; and
CC treating cancer in a mammal by administering an effective amount of a
CC human anti-CTLA-4 antibody to a mammal who has undergone stem cell
CC transplantation. The cancer is chosen from breast cancer including
CC metastatic breast cancer, lung cancer including small-cell lung cancer,
CC bone cancer, pancreatic cancer, skin cancer, cancer of the head or neck,
CC melanoma including coetaneous or interlobular malignant melanoma, uterine
CC cancer, ovarian cancer, rectal cancer, cancer of the anal region, stomach
CC cancer, colon cancer, testicular cancer, uterine cancer, carcinoma of the
CC fallopian tubes, carcinoma of the endometrium, carcinoma of the cervix,
CC carcinoma of the vagina, carcinoma of the vulva, Hodgkin's disease, non-
CC Hodgkin's lymphoma, cancer of the esophagus, cancer of the small
CC intestine, cancer of the endocrine system, cancer of the thyroid gland,
CC cancer of the parathyroid gland, cancer of the adrenal gland, sarcoma of
CC soft tissue, cancer of the urethra, cancer of the penis, prostate cancer,
CC chronic acute leukemias including acute myeloid leukemia, chronic myeloid

CC leukemia, acute lymphoblastic leukemia, chronic lymphocytic leukemia,
CC solid tumors of childhood, lymphocytic lymphomas, cutaneous T cell
CC lymphoma, cancer of the bladder, cancer of the kidney or ureter, renal
CC cell carcinoma, carcinoma of the renal pelvis, neoplasm of the central
CC nervous system (CNS), primary CNS lymphoma, tumor angiogenesis, spinal
CC axis tumor, brain stem glioma, pituitary adenoma, Kaposi's sarcoma,
CC epidermoid cancer, squamous cell cancer, T-cell lymphoma, environmentally
CC induced cancers including those induced by asbestos, myeloma,
CC neuroblastoma, and pediatric sarcomas. The present sequence represents
CC human anti-CTLA-4 antibody, kappa chain protein SEQ ID NO:36.

XX
SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQRPQGSPPRLIYKVSNRDSGVPDRFSGSGGTFTLKISRVEAED 220
|||||
DB 40 WFQRPQGSPPRLIYKVSNRDSGVPDRFSGSGGTFTLKISRVEAED 87
|||||

RESULT 15
ID AED06937
XX AED06937 standard; protein; 100 AA.
AC AED06937;
XX
DT 01-DEC-2005 (first entry)
XX
DE Human antibody V kappa protein sequence, SEQ ID NO:107.
XX
KW antibody therapy; antibody engineering; transgenic animal;
KW drug screening; leukemia; hematological disease; neoplasm; cytostatic;
KW autoimmune disease; immunosuppressive; immune disorder.
XX
OS Homo sapiens.
XX
XX WO2005092926-A2.
PN
XX
PD 06-OCT-2005.
XX
XX 17-MAR-2005; 2005WO-US009306.
XX
XX 19-MAR-2004; 2004US-0554372P.
PR
XX 24-MAY-2004; 2004US-0574661P.
XX
XX (ABGE-) ABGENIX INC.
XX
XX Kellermann S, Green LL, Korver W;
XX
XX WPI; 2005-664987/68.
XX
PT Selecting an antibody with a reduced risk of inducing a human anti-human
PT antibody response comprises providing an immunoglobulin gene encoding a
PT candidate antibody.
XX
XX Disclosure; SEQ ID NO 107; 235pp; English.

XX The new invention relates to selecting an antibody with a reduced risk of
CC inducing an anti-human antibody (HAHA) response in the host. The method
CC comprises providing an immunoglobulin gene encoding a candidate antibody
CC and comparing it with an immunoglobulin gene from the recipient. Also
CC claimed is a method of excluding an antibody from use in the treatment of
CC a host. Also given are transgenic animals and methods for identifying
CC antibodies or for screening or agents that inhibit the induction of HAHA
CC response; kit for detecting an antibody or assessing the risk of a HAHA
CC response; a method of determining a risk that an antibody will induce
CC HAHA response in a patient; and a method of increasing the probability
CC that a HAHA response will be detected in a transgenic mouse. The method
CC further comprises repeating the steps of providing, comparing, and
CC selecting for more than one immunoglobulin gene or for every
CC immunoglobulin V gene of the candidate antibody. The immunoglobulin gene

CC is a V gene, e.g. a VH (heavy) gene or a VL (light) gene. The methods are
CC useful for selecting an antibody with a reduced risk of inducing HAMA
CC response. The antibody, composition, and method are useful for treating
CC diseases, e.g. leukemia or autoimmune disease. The present sequence is a
CC human antibody V kappa protein sequence.
XX

SQ Sequence 100 AA;

Query Match 18.4%; Score 48; DB 9; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.5e-32;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGYDPDRFSGSGGTDFTLKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVSNRDSGYDPDRFSGSGGTDFTLKISRVEAD 87

Search completed: August 9, 2006, 13:34:15
Job time : 196 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:39:38 ; Search time 49 Seconds
(without alignments)
466.235 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVYKDDDK 261

Scoring table:

OLIGO 60.0 , Gapext 60.0

Searched: 650591 seqs, 87530628 residues

Word size: 1

Total number of hits satisfying chosen parameters: 649417

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Issued Patents AA.*

- 1: /EMC Celleria SIDS3/ptodata/2/iaa/5_COMB.pap.*
- 2: /EMC Celleria SIDS3/ptodata/2/iaa/6_COMB.pap.*
- 3: /EMC Celleria SIDS3/ptodata/2/iaa/7_COMB.pap.*
- 4: /EMC Celleria SIDS3/ptodata/2/iaa/H_COMB.pap.*
- 5: /EMC Celleria SIDS3/ptodata/2/iaa/PCTUS_COMB.pap.*
- 6: /EMC Celleria SIDS3/ptodata/2/iaa/RE_COMB.pap.*
- 7: /EMC Celleria SIDS3/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	48	18.4	100	2 US-09-472-087-113 Sequence 113, App
2	48	18.4	100	2 US-10-194-975-75 Sequence 75, App
3	48	18.4	111	2 US-09-809-739-13 Sequence 13, App
4	48	18.4	111	2 US-09-840-459-11 Sequence 11, App
5	48	18.4	111	2 US-09-840-459-59 Sequence 59, App
6	48	18.4	111	2 US-09-497-625A-11 Sequence 11, App
7	48	18.4	111	2 US-09-497-625A-59 Sequence 59, App
8	48	18.4	112	1 US-07-942-245-28 Sequence 28, App
9	48	18.4	112	1 US-08-477-877B-89 Sequence 89, App
10	48	18.4	112	1 US-08-472-281A-89 Sequence 89, App
11	48	18.4	112	1 US-08-477-989B-89 Sequence 89, App
12	48	18.4	112	2 US-09-840-459-56 Sequence 56, App
13	48	18.4	112	2 US-09-840-459-62 Sequence 62, App
14	48	18.4	112	2 US-09-840-459-66 Sequence 66, App
15	48	18.4	112	2 US-09-497-625A-56 Sequence 56, App
16	48	18.4	112	2 US-09-497-625A-62 Sequence 62, App
17	48	18.4	112	2 US-09-497-625A-66 Sequence 66, App
18	48	18.4	112	2 US-09-254-180C-8 Sequence 8, App
19	48	18.4	112	2 US-09-462-140D-97 Sequence 97, App
20	46	17.6	112	2 US-09-840-459-60 Sequence 60, App
21	46	17.6	112	2 US-09-497-625A-60 Sequence 60, App
22	37	14.2	112	2 US-09-840-459-70 Sequence 70, App
23	37	14.2	112	2 US-09-497-625A-70 Sequence 70, App
24	34	13.0	96	2 US-10-330-613A-54 Sequence 54, App
25	34	13.0	98	2 US-10-194-975-4 Sequence 4, App
26	34	13.0	98	2 US-10-330-613A-53 Sequence 53, App

27 34 13.0 112 2 US-09-840-459-69 Sequence 69, Appl
28 34 13.0 112 2 US-09-497-625A-69 Sequence 69, Appl
29 34 13.0 117 2 US-08-545-809A-105 Sequence 105, App
30 34 13.0 117 2 US-09-515-697-105 Sequence 105, App
31 34 13.0 123 2 US-10-330-613A-21 Sequence 21, Appl
32 34 13.0 288 2 US-09-818-247-22 Sequence 22, Appl
33 34 12.6 112 2 US-09-840-459-67 Sequence 67, Appl
34 34 12.6 112 2 US-09-497-625A-67 Sequence 67, Appl
35 32 12.3 112 2 US-09-840-459-61 Sequence 61, Appl
36 32 12.3 112 2 US-09-497-625A-61 Sequence 61, Appl
37 32 12.3 113 2 US-09-840-459-63 Sequence 63, Appl
38 32 12.3 113 2 US-09-497-625A-63 Sequence 63, Appl
39 31 11.9 138 2 US-09-453-718B-90 Sequence 90, Appl
40 30 11.5 88 2 US-09-254-180C-151 Sequence 151, App
41 30 11.5 112 2 US-09-840-459-58 Sequence 58, Appl
42 30 11.5 112 1 US-08-428-257A-78 Sequence 78, Appl
43 30 11.5 118 2 US-09-199-149-5 Sequence 5, Appl
44 30 11.5 118 2 US-07-987-264-60 Sequence 60, Appl
45 30 11.5 118 2 US-07-987-264-60 Sequence 60, Appl

ALIGNMENTS

RESULT 1
US-09-472-087-113
; Sequence 113, Application US/09472087
; Patent No. 6682736
; GENERAL INFORMATION:
; APPLICANT: HANSON, DOUGLAS C.
; APPLICANT: NEVEU, MARK J.
; APPLICANT: MUELLER, EILLEN E.
; APPLICANT: HANKE, JEFFREY H.
; APPLICANT: GILMAN, STEVEN C.
; APPLICANT: DAVIS, C. GEORFREY
; APPLICANT: CORVALAN, JOSE R.
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PFI
; CURRENT APPLICATION NUMBER: US/09/472,087
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-472-087-113
Query Match 18.4%; Score 48; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.8e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFQRPQSGSPRLIYKVSNRDGVDPFSGSGGTDTTLKISRVEAD 220
Db 40 WFQRPQSGSPRLIYKVSNRDGVDPFSGSGGTDTTLKISRVEAD 87
RESULT 2
US-10-194-975-75
; Sequence 75, Application US/10194975
; Patent No. 6881557
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-75

Query Match      18.4%; Score 48; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.8e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 220
    |||||
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 87

RESULT 3
US-09-809-739-13
; Sequence 13, Application US/09809739
; Patent No. 6663863
; GENERAL INFORMATION:
; APPLICANT: Horvath, Christopher J.
; APPLICANT: Rao, Patricia E.
; TITLE OF INVENTION: Method of Inhibiting Stenosis and
; TITLE OF INVENTION: Restenosis
; FILE REFERENCE: 1855.1069-003
; CURRENT APPLICATION NUMBER: US/09/809,739
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: US 09/528,267
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-739-13

Query Match      18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 220
    |||||
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 87

RESULT 4
US-09-840-459-11
; Sequence 11, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR FILING DATE: 2000-02-03
; PRIOR FILING DATE: 1999-07-22
; PRIOR FILING DATE: 09/121,781
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-11

Query Match      18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 220
    |||||
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 87
```

```
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-11

Query Match      18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 220
    |||||
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 87

RESULT 5
US-09-840-459-59
; Sequence 59, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-59

Query Match      18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 220
    |||||
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPDRFSGSGGTDTFTLKISRVEAD 87

RESULT 6
US-09-497-625A-11
; Sequence 11, Application US/09497625A
; Patent No. 6727349
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-004
; CURRENT APPLICATION NUMBER: US/09/497,625A
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
```

; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-497-625A-11

Query Match 18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPGQSPRLIYKVNRSQVDPDRFSGSGTDTLTKISRVEAD 220
|||||
Db 40 WFOQRPGQSPRLIYKVNRSQVDPDRFSGSGTDTLTKISRVEAD 87

RESULT 7

US-09-497-625A-59
; Sequence 59, Application US/09497625A
; Patent No. 6727349
; GENERAL INFORMATION:

; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa

; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; FILE REFERENCE: 1855.1052-004

; CURRENT APPLICATION NUMBER: US/09/497,625A

; PRIOR FILING DATE: 2000-02-03

; PRIOR FILING DATE: 1999-07-22

; PRIOR APPLICATION NUMBER: 09/121,781

; PRIOR FILING DATE: 1998-07-23

; NUMBER OF SEQ ID NOS: 106

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 59

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-497-625A-59

Query Match 18.4%; Score 48; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 8.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPGQSPRLIYKVNRSQVDPDRFSGSGTDTLTKISRVEAD 220
|||||
Db 40 WFOQRPGQSPRLIYKVNRSQVDPDRFSGSGTDTLTKISRVEAD 87

RESULT 8

US-07-942-245-28
; Sequence 28, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:

; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.

; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
; NUMBER OF SEQUENCES: 522

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas

; STREET: 2100 Pennsylvania Avenue, N.W.

; CITY: Washington

; STATE: D. C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: HP 9000/700 Workstation
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-28

Query Match 18.4%; Score 48; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPGQSPRLIYKVNRSQVDPDRFSGSGTDTLTKISRVEAD 220
|||||
Db 40 WFOQRPGQSPRLIYKVNRSQVDPDRFSGSGTDTLTKISRVEAD 87

RESULT 9

US-08-477-877B-89

; Sequence 89, Application US/08477877B

; Patent No. 5730979

; GENERAL INFORMATION:

; APPLICANT: Bazin, Herv

; APPLICANT: Latanne, Dominique

; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation

; NUMBER OF SEQUENCES: 96

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carella, Byrnie, Bain, Gilfillan,

; ADDRESSEE: Cecchi, Stewart & Olstein

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: New Jersey

; COUNTRY: U.S.A.

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,877B

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/407,009

; FILING DATE: 29-MAR-1995

; APPLICATION NUMBER: 08/119,032

; FILING DATE: 09-SEP-1993

; APPLICATION NUMBER: 08/027,008

; FILING DATE: 05-MAR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Olstein, Elliot M.

; REGISTRATION NUMBER: 24,025

; REFERENCE/DOCKET NUMBER: 61750-146

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

```

; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Light chain vairable region of HUM5400
; US-08-477-877B-89

Query Match      18.4%; Score 48; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFLKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFLKISRVEAD 87

RESULT 10
US-08-472-281A-89
; Sequence 89, Application US/08472281A
; Patent No. 5817311
; GENERAL INFORMATION:
; APPLICANT: Bazin, Herv
; APPLICANT: Latinne, Dominique
; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.281A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,009
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: 08/119,032
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: 08/027,008
; FILING DATE: 05-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 61750-142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Light chain vairable region of HUM5400
; US-08-472-281A-89

Query Match      18.4%; Score 48; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFLKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFLKISRVEAD 87

RESULT 11
US-08-477-989B-89
; Sequence 89, Application US/08477989B
; Patent No. 5951983
; GENERAL INFORMATION:
; APPLICANT: Bazin, Herv
; APPLICANT: Latinne, Dominique
; APPLICANT: Kaplan, Ruth
; APPLICANT: Kieber-Emmons, Thomas
; APPLICANT: Postema, Christina E.
; APPLICANT: White-Scharf, Mary
; TITLE OF INVENTION: LO-CD2a Antibody and Uses
; TITLE OF INVENTION: Thereof for Inhibiting
; TITLE OF INVENTION: T-Cell Activation and
; TITLE OF INVENTION: Proliferation
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,989B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,009
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: 08/119,032
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: 08/027,008
; FILING DATE: 05-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 61750-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Light chain vairable region of
; NAME/KEY: HUM5400
; US-08-477-989B-89

Query Match      18.4%; Score 48; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFLKISRVEAD 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFLKISRVEAD 87
```

RESULT 12
US-09-840-459-56
; Sequence 56, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-56

Query Match 18.4%; Score 48; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPRLIYKVNSRDGVPDRFSGSGGTDFTLKISRVEAD 220
|||
Db 40 WFOQPGQSPRLIYKVNSRDGVPDRFSGSGGTDFTLKISRVEAD 87

RESULT 13
US-09-840-459-62
; Sequence 62, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-840-459-62

Query Match 18.4%; Score 48; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPRLIYKVNSRDGVPDRFSGSGGTDFTLKISRVEAD 220
|||
Db 40 WFOQPGQSPRLIYKVNSRDGVPDRFSGSGGTDFTLKISRVEAD 87

RESULT 14
US-09-840-459-66
; Sequence 66, Application US/09840459
; Patent No. 6696550
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-012
; CURRENT APPLICATION NUMBER: US/09/840,459
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: PCT/US01/03537
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/497,625
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-459-66

Query Match 18.4%; Score 48; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGQSPRLIYKVNSRDGVPDRFSGSGGTDFTLKISRVEAD 220
|||
Db 40 WFOQPGQSPRLIYKVNSRDGVPDRFSGSGGTDFTLKISRVEAD 87

RESULT 15
US-09-497-625A-56
; Sequence 56, Application US/09497625A
; Patent No. 6727349
; GENERAL INFORMATION:
; APPLICANT: LaRosa, Gregory J.
; APPLICANT: Horvath, Christopher
; APPLICANT: Newman, Walter
; APPLICANT: Jones, S. Tarran
; APPLICANT: O'Brien, Siobhan H.
; APPLICANT: O'Keefe, Theresa
; TITLE OF INVENTION: HUMANIZED ANTI-CCR2 ANTIBODIES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 1855.1052-004
; CURRENT APPLICATION NUMBER: US/09/497,625A
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/359,193
; PRIOR FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: 09/121,781
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 106

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 56

; LENGTH: 112

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-497-625A-56

Query Match 18.4%; Score 48; DB 2; Length 112;

Best Local Similarity 100.0%; Pred.No. 8.5e-30;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFOQPGOSPRRLIYKVSNRDVGVPDRFSGSGGTDFTLKISRVEAED 220

Db 40 WFOQPGOSPRRLIYKVSNRDVGVPDRFSGSGGTDFTLKISRVEAED 87

Search completed: August 9, 2006, 13:41:02

Job time : 50 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:51:33 ; Search time 176 Seconds
(without alignments)
686.927 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2097797 seqs, 463214858 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2096646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Published Applications AA Main.*
- 1: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 3: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 4: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 5: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 6: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261	100.0	261	6	US-11-056-825-2
2	217	83.1	259	6	US-11-056-825-7
3	48	18.4	100	4	US-10-194-975-75
4	48	18.4	100	4	US-10-153-382-36
5	48	18.4	100	4	US-10-308-817-22
6	48	18.4	100	4	US-10-453-698-22
7	48	18.4	100	4	US-10-379-392-87
8	48	18.4	100	5	US-10-612-497-113
9	48	18.4	100	5	US-10-776-649-113
10	48	18.4	100	6	US-11-085-368-36
11	48	18.4	100	6	US-11-084-554-107
12	48	18.4	100	6	US-11-054-669-75
13	48	18.4	100	6	US-11-128-900-113
14	48	18.4	100	6	US-11-004-590-82
15	48	18.4	100	6	US-11-136-250-107
16	48	18.4	111	3	US-09-835-087-2
17	48	18.4	111	3	US-09-809-739-13
18	48	18.4	111	3	US-09-840-459-11
19	48	18.4	111	3	US-09-840-459-59
20	48	18.4	111	4	US-10-766-773-11
21	48	18.4	111	4	US-10-766-773-59
22	48	18.4	111	4	US-10-766-610-11
23	48	18.4	111	4	US-10-766-610-59
24	48	18.4	111	4	US-10-733-563-11
25	48	18.4	111	4	US-10-733-563-59
26	48	18.4	111	5	US-10-706-852-15
27	48	18.4	111	5	US-10-662-061-13

28	48	18.4	112	3	US-09-840-459-56	Sequence 56, Appl
29	48	18.4	112	3	US-09-840-459-62	Sequence 62, Appl
30	48	18.4	112	3	US-09-840-459-66	Sequence 66, Appl
31	48	18.4	112	4	US-10-766-773-56	Sequence 56, Appl
32	48	18.4	112	4	US-10-766-773-62	Sequence 62, Appl
33	48	18.4	112	4	US-10-766-773-66	Sequence 66, Appl
34	48	18.4	112	4	US-10-766-610-56	Sequence 56, Appl
35	48	18.4	112	4	US-10-766-610-62	Sequence 62, Appl
36	48	18.4	112	4	US-10-766-610-66	Sequence 66, Appl
37	48	18.4	112	4	US-10-733-563-56	Sequence 56, Appl
38	48	18.4	112	4	US-10-733-563-62	Sequence 62, Appl
39	48	18.4	112	4	US-10-733-563-66	Sequence 66, Appl
40	48	18.4	253	3	US-09-880-748-1964	Sequence 1964, Ap
41	48	18.4	253	4	US-10-293-418-1964	Sequence 1964, Ap
42	48	18.4	253	6	US-11-054-515-1964	Sequence 1964, Ap
43	48	18.4	253	6	US-11-266-444-1964	Sequence 1964, Ap
44	46	17.6	112	3	US-09-840-459-60	Sequence 60, Appl
45	46	17.6	112	4	US-10-766-773-60	Sequence 60, Appl

ALIGNMENTS

RESULT 1

US-11-056-825-2
; Sequence 2, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Janda, Kim D.
; APPLICANT: Janda, Kim D.
; APPLICANT: Janda, Kim D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-2

Query Match		100.0%;	Score 261;	DB 6;	Length 261;
Best Local Similarity		100.0%;	Pred. No. 7.4e-201;		
Matches 261;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAQVLVQSGAEVKKPGASVKVSKASGYTPNIGITVVRQAPQGLWGWINGNTHY	60		
Db	1	MAQVLVQSGAEVKKPGASVKVSKASGYTFNIGITVVRQAPQGLWGWINGNTHY	60		
Qy	61	AKQFQGRVTMTDTSTSTAYMELRSLSDDTAVYVCARDPRGDDPEYWGOGTLVTYSSGG	120		
Db	61	AKQFQGRVTMTDTSTSTAYMELRSLSDDTAVYVCARDPRGDDPEYWGOGTLVTYSSGG	120		
Qy	121	GGSGGGGGGGGGSEIVLTQSPSLPVLTLQGPASISCRSSQNLVYSDGNTVLSWFQORPGQ	180		
Db	121	GGSGGGGGGGGGSEIVLTQSPSLPVLTLQGPASISCRSSQNLVYSDGNTVLSWFQORPGQ	180		
Qy	181	SPRRLIYKVNDRSGVDPFRFGSGSGTDTFTLKISRVEAEDIGVYCMQGTWHPRTFGQ	240		
Db	181	SPRRLIYKVNDRSGVDPFRFGSGSGTDTFTLKISRVEAEDIGVYCMQGTWHPRTFGQ	240		
Qy	241	TKVEIKRGLGLVDYKDDDK	261		
Db	241	TKVEIKRGLGLVDYKDDDK	261		

RESULT 2
US-11-056-825-7
; Sequence 7, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SCRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: US 60/544,807
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-7
Query Match 83.1%; Score 217; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.3e-165; Indels 0; Gaps 0;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTYVRRQAPGGLEWMGWINNNTYH 60
Db 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTYVRRQAPGGLEWMGWINNNTYH 60
Qy 61 AQKPGQRTVMTDTSTSTAYNELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 AQKPGQRTVMTDTSTSTAYNELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Qy 121 GSGGGSGGGGSEIVLTQSPLSIPLVTIGQPASISCRSQNLVYSDGNTYLSWFOORPCQ 180
Db 121 GSGGGSGGGGSEIVLTQSPLSIPLVTIGQPASISCRSQNLVYSDGNTYLSWFOORPCQ 180
Qy 181 SPRLLIYKVSNRDGVDPDFSGSGGTDTLTKISRVE 217
Db 181 SPRLLIYKVSNRDGVDPDFSGSGGTDTLTKISRVE 217
RESULT 3
US-10-194-975-75
; Sequence 75, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-75
Query Match 18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPGQSPRLIYKVSNRDGVDPDFSGSGGTDTLTKISRVEAD 220

Db 40 WFOQRPGQSPRLIYKVSNRDGVDPDFSGSGGTDTLTKISRVEAD 87
RESULT 4
US-10-153-382-36
; Sequence 36, Application US/10153382
; Publication No. US20030086930A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER PRODUCTS INC.
; TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
; FILE REFERENCE: PC23019A
; CURRENT APPLICATION NUMBER: US/10/153,382
; CURRENT FILING DATE: 2002-05-22
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-382-36
Query Match 18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPGQSPRLIYKVSNRDGVDPDFSGSGGTDTLTKISRVEAD 220
Db 40 WFOQRPGQSPRLIYKVSNRDGVDPDFSGSGGTDTLTKISRVEAD 87
RESULT 5
US-10-308-817-22
; Sequence 22, Application US/10308817
; Publication No. US20030219861A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; APPLICANT: Wu, Davang
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
US-10-308-817-22
Query Match 18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPGQSPRLIYKVSNRDGVDPDFSGSGGTDTLTKISRVEAD 220
Db 40 WFOQRPGQSPRLIYKVSNRDGVDPDFSGSGGTDTLTKISRVEAD 87
RESULT 6
US-10-453-698-22
; Sequence 22, Application US/10453698
; Publication No. US20040038308A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 82 CIP (1087-37 CIP)
; CURRENT APPLICATION NUMBER: US/10/453,698
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.2

```
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
; US-10-453-698-22

Query Match      18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 7
US-10-379-392-87
; Sequence 87, Application US/10379392
; Publication No. US20040110226A1
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahivat, Bassil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: A-71386-3 463077-236
; CURRENT APPLICATION NUMBER: US/10/379,392
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-379-392-87

Query Match      18.4%; Score 48; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 8
US-10-612-497-113
; Sequence 113, Application US/10612497
; Publication No. US20040228859A1
; GENERAL INFORMATION:
; APPLICANT: Douglas C. Hanson
; APPLICANT: Mark J. Neveu
; APPLICANT: Eileen E. Mueller
; APPLICANT: Jeffrey H. Hanke
; APPLICANT: Steven C. Gilman
; APPLICANT: C. Geoffrey Davis
; APPLICANT: Jose R. Corvalan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PF1 DIV1
; CURRENT APPLICATION NUMBER: US/10/612,497
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-612-497-113

Query Match      18.4%; Score 48; DB 5; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 9
US-10-776-649-113
; Sequence 113, Application US/10776649
; Publication No. US20040228861A1
; GENERAL INFORMATION:
; APPLICANT: Douglas C. Hanson
; APPLICANT: Mark J. Neveu
; APPLICANT: Eileen E. Mueller
; APPLICANT: Jeffrey H. Hanke
; APPLICANT: Steven C. Gilman
; APPLICANT: C. Geoffrey Davis
; APPLICANT: Jose R. Corvalan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PF1 DIV2
; CURRENT APPLICATION NUMBER: US/10/776,649
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 10/612497
; PRIOR FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-776-649-113

Query Match      18.4%; Score 48; DB 5; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVSNRDSGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 10
US-11-085-368-36
; Sequence 36, Application US/11085368
; Publication No. US20050226875A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER PRODUCTS INC.
; APPLICANT: Gomez-Navarro, Jesus
; APPLICANT: Hanson, Douglas C.
; APPLICANT: Eileen, Mueller Elliott
; APPLICANT: Noe, Dennis A.
; TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
; FILE REFERENCE: PC32177A
; CURRENT APPLICATION NUMBER: US/11/085,368
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/556,801
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 100
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-11-085-368-36

Query Match      18.4%; Score 48; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 87

RESULT 11
US-11-084-554-107
; Sequence 107, Application US/11084554
; Publication No. US20050260679A1
; GENERAL INFORMATION:
; APPLICANT: Kellermann, Sirid-Ai
; APPLICANT: Green, Larry L.
; APPLICANT: Korvet, Wouter
; TITLE OF INVENTION: REDUCING THE RISK OF HUMAN ANTI-HUMAN
; TITLE OF INVENTION: ANTIBODIES THROUGH V GENE MANIPULATION
; FILE REFERENCE: ABGENIX 100A
; CURRENT APPLICATION NUMBER: US/11/084,554
; CURRENT FILING DATE: 2005-03-17
; PRIOR APPLICATION NUMBER: 60/554,372
; PRIOR FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/574,661
; PRIOR FILING DATE: 2004-05-24
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-084-554-107

Query Match      18.4%; Score 48; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 87

RESULT 12
US-11-054-669-75
; Sequence 75, Application US/11054669
; Publication No. US20050261480A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: SUPER HUMANIZED ANTIBODIES
; FILE REFERENCE: 30219/US/3
; CURRENT APPLICATION NUMBER: US/11/054,669
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 10/194,975
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-054-669-75

Query Match      18.4%; Score 48; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 87

US-11-056-825-2.olig.rapbm

Db 40 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 87

RESULT 13
US-11-128-900-113
; Sequence 113, Application US/11128900
; Publication No. US20050287136A1
; GENERAL INFORMATION:
; APPLICANT: HANSON, DOUGLAS C.
; APPLICANT: NEVEU, MARK J.
; APPLICANT: MUELLER, EILEEN E.
; APPLICANT: HANKE, JEFFREY H.
; APPLICANT: GILMAN, STEVEN C.
; APPLICANT: DAVIS, C. GEOFFREY
; APPLICANT: CORVALAN, JOSE R.
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PFI DIV3
; CURRENT APPLICATION NUMBER: US/11/128,900
; CURRENT FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: US 10/776649
; PRIOR FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 10/612497
; PRIOR APPLICATION NUMBER: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-128-900-113

Query Match      18.4%; Score 48; DB 6; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-30;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 220
Db 40 WFOQRPQSPRLIYKVNSRDSGVDPFRFSGSGTDFTLKISRVEAED 87

RESULT 14
US-11-004-590-82
; Sequence 82, Application US/11004590
; Publication No. US2006000883A1
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John R.
; APPLICANT: Hammond, Phillip W.
; TITLE OF INVENTION: METHODS OF GENERATING VARIANT PROTEINS WITH INCREASED HOST STRING
; TITLE OF INVENTION: CONTENT AND COMPOSITIONS THEREOF
; FILE REFERENCE: 185832/US/5
; CURRENT APPLICATION NUMBER: US/11/004,590
; CURRENT FILING DATE: 2004-12-03
; PRIOR APPLICATION NUMBER: US 60/527,167
; PRIOR FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: US 60/581,613
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 60/601,665
; PRIOR FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US 60/619,483
; PRIOR FILING DATE: 2004-10-14
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-004-590-82
```

Search completed: August 9, 2006, 13:54:56
Job time : 177 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:52:08 ; Search time 33 Seconds ;
(without alignments)
532.377 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVYKDDDDK 261

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 239914 seqs, 67312017 residues

Word size : 1

Total number of hits satisfying chosen parameters: 239854

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA New:

- 1: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	18.4	100	7	US-11-239-308-45
2	48	18.4	113	7	US-11-239-308-2
3	35	13.4	247	6	US-10-539-402-18
4	34	13.0	98	7	US-11-221-902-55
5	33	12.6	118	7	US-11-375-221-11
6	31	11.9	32	6	US-10-497-088-7
7	31	11.9	120	7	US-11-337-300-33
8	31	11.9	247	7	US-11-337-300-65
9	31	11.9	291	7	US-11-154-103-10
10	31	11.9	1052	6	US-10-497-088-21
11	31	11.9	1342	6	US-10-497-088-14
12	30	11.5	117	7	US-11-249-296-6
13	30	11.5	453	7	US-11-254-182-44
14	29	11.1	30	6	US-10-570-220-49
15	29	11.1	30	7	US-11-221-902-49
16	29	11.1	87	7	US-11-219-121-19
17	29	11.1	87	7	US-11-061-841-23
18	29	11.1	87	7	US-11-196-917A-19
19	29	11.1	98	7	US-11-221-902-52
20	29	11.1	98	7	US-11-221-902-53
21	29	11.1	98	7	US-11-221-902-54
22	29	11.1	98	7	US-11-221-902-58
23	29	11.1	98	7	US-11-221-902-64
24	29	11.1	108	6	US-10-484-105-14
25	29	11.1	118	7	US-11-249-296-2

26	29	11.1	118	7	US-11-249-296-72	Sequence 72, Appl
27	29	11.1	118	7	US-11-249-296-90	Sequence 90, Appl
28	29	11.1	119	7	US-11-006-808-10	Sequence 10, Appl
29	29	11.1	119	7	US-11-291-140-25	Sequence 25, Appl
30	29	11.1	119	7	US-11-291-140-27	Sequence 27, Appl
31	29	11.1	119	7	US-11-291-140-29	Sequence 29, Appl
32	29	11.1	119	7	US-11-291-140-45	Sequence 45, Appl
33	29	11.1	119	7	US-11-291-140-47	Sequence 47, Appl
34	29	11.1	119	7	US-11-291-140-49	Sequence 49, Appl
35	29	11.1	119	7	US-11-291-140-51	Sequence 51, Appl
36	29	11.1	119	7	US-11-291-140-53	Sequence 53, Appl
37	29	11.1	120	7	US-11-304-986-22	Sequence 22, Appl
38	29	11.1	122	7	US-11-211-917-110	Sequence 110, Appl
39	29	11.1	126	7	US-11-211-917-42	Sequence 42, Appl
40	29	11.1	129	7	US-11-006-808-45	Sequence 45, Appl
41	29	11.1	139	6	US-10-533-104A-21	Sequence 21, Appl
42	29	11.1	139	6	US-10-533-104A-22	Sequence 22, Appl
43	29	11.1	247	7	US-11-337-300-96	Sequence 96, Appl
44	29	11.1	248	7	US-11-337-300-94	Sequence 94, Appl
45	29	11.1	249	7	US-11-337-300-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1

US-11-239-308-45
; Sequence 45, Application US/11239308
; Publication No. US2006008883A1
; GENERAL INFORMATION:
; APPLICANT: Snider, Vaughn
; APPLICANT: Larrick, James W.
; APPLICANT: Integrigen, Inc.
; TITLE OF INVENTION: Recombinant Catalytic Polypeptides and Their Uses
; FILE REFERENCE: 021216-000310US
; CURRENT APPLICATION NUMBER: US/11/239,308
; PRIOR FILING DATE: 2005-09-28
; PRIOR APPLICATION NUMBER: US/10/683,733
; PRIOR FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US 60/417,979
; PRIOR FILING DATE: 2002-10-09
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 45
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-239-308-45

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 8.6e-33;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQRPQSGPRRLIYKVNRSRDSVDFPFGSGSGTFTLKISRVEARD 220
DB 40 WFQRPQSGPRRLIYKVNRSRDSVDFPFGSGSGTFTLKISRVEARD 87

RESULT 2

US-11-239-308-2
; Sequence 2, Application US/11239308
; Publication No. US2006008883A1
; GENERAL INFORMATION:
; APPLICANT: Snider, Vaughn
; APPLICANT: Larrick, James W.
; APPLICANT: Integrigen, Inc.
; TITLE OF INVENTION: Recombinant Catalytic Polypeptides and Their Uses
; FILE REFERENCE: 021216-000310US
; CURRENT APPLICATION NUMBER: US/11/239,308
; PRIOR FILING DATE: 2005-09-28
; PRIOR APPLICATION NUMBER: US/10/683,733
; PRIOR FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US 60/417,979

; PRIOR FILING DATE: 2002-10-09
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-239-308-2

Query Match 18.4%; Score 48; DB 7; Length 113;
Best Local Similarity 100.0%; Pred. No. 9.4e-33;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOQPQSPRRLIYKVSNRDVGVPDRFSGSGGTDTFLKISRVEAED 220
Db 40 WFOQPQSPRRLIYKVSNRDVGVPDRFSGSGGTDTFLKISRVEAED 87

RESULT 3
US-10-539-402-18
; Sequence 18, Application US/10539402
; Publication No. US20060115477A1
; GENERAL INFORMATION:
; APPLICANT: Xerion Pharmaceuticals AG
; TITLE OF INVENTION: Neuropilin-1 Inhibitor
; FILE REFERENCE: XE12EPC
; CURRENT APPLICATION NUMBER: US/10/539,402
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US 60/435,893
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: EP 03000615
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 247
; TYPE: PRT
; ORGANISM: human
US-10-539-402-18

Query Match 13.4%; Score 35; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 YAKQFQGRVTMTTDTSTSTAYMELRSLRSDDTAVY 94
Db 52 YAKQFQGRVTMTTDTSTSTAYMELRSLRSDDTAVY 86

RESULT 4
US-11-221-902-55
; Sequence 55, Application US/11221902
; Publication No. US20060088522A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: HUMANIZED ANTI-5T4 ANTIBODIES AND ANTI-5T4/CALICHEAMICIN CONJUGATE
; FILE REFERENCE: 040000-0317285
; CURRENT APPLICATION NUMBER: US/11/221,902
; CURRENT FILING DATE: 2005-09-09
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-221-902-55

Query Match 13.0%; Score 34; DB 7; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.6e-21;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 QGRVTMTTDTSTSTAYMELRSLRSDDTAVYCAR 98

Db 65 QGRVTMTTDTSTSTAYMELRSLRSDDTAVYCAR 98

RESULT 5
US-11-375-221-111
; Sequence 111, Application US/11375221
; Publication No. US20060153850A1
; GENERAL INFORMATION:
; APPLICANT: BEDIAN, VAHE
; APPLICANT: DEVALARAJA, MADHAV NARASIMHA
; APPLICANT: FOLTZ, IAN
; APPLICANT: HAAK-FRENSCH, MARY
; APPLICANT: KELLERMANN, SIRID-AIMEE
; APPLICANT: LOW, JOSEPH EDWIN
; APPLICANT: MOBLEY, JAMES LESLIE
; TITLE OF INVENTION: ANTIBODIES TO M-CSF
; FILE REFERENCE: ABX-PP4
; CURRENT APPLICATION NUMBER: US/11/375,221
; CURRENT FILING DATE: 2006-03-13
; PRIOR APPLICATION NUMBER: US/10/938,353
; PRIOR FILING DATE: 2004-09-09
; PRIOR APPLICATION NUMBER: 60/502,163
; PRIOR FILING DATE: 2003-09-10
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 111
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-375-221-111

Query Match 12.6%; Score 33; DB 7; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.8e-20;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 QGRVTMTTDTSTSTAYMELRSLRSDDTAVYCAR 97
Db 65 QGRVTMTTDTSTSTAYMELRSLRSDDTAVYCAR 97

RESULT 6
US-10-497-088-7
; Sequence 7, Application US/10497088
; Publication No. US20060088520A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: Germeaad, Wilfred
; APPLICANT: Logtenberg, Ton
; APPLICANT: Lekkerkerker, Annemarie N
; TITLE OF INVENTION: Antigen presenting cell targeting conjugate, an antigen
; TITLE OF INVENTION: presenting cell contacted with such conjugate, their use for
; TITLE OF INVENTION: vaccination or as medicament, and methods for their production or
; TITLE OF INVENTION: generation
; FILE REFERENCE: 0070 US 00 CON
; CURRENT APPLICATION NUMBER: US/10/497,088
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: PCT/EP01/14255
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/EP02/13681
; PRIOR FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: EP01204997.9
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: FRI of MatDC16
US-10-497-088-7

Query Match 11.9%; Score 31; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 5e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31
Db 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31

RESULT 7

US-11-337-300-33
; Sequence 33, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruijf, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337.300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Variable heavy chain of SC03-012
US-11-337-300-33

Query Match 11.9%; Score 31; DB 7; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31
Db 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31

RESULT 8

US-11-337-300-65
; Sequence 65, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruijf, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337.300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: SC03-012
US-11-337-300-65

Query Match 11.9%; Score 31; DB 7; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.2e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 31
Db 2 MAQVQLVQSGAEVKKPGASVKVSKASGYTF 32

RESULT 9

US-11-154-103-10
; Sequence 10, Application US/11154103
; Publication No. US20060099205A1
; GENERAL INFORMATION:
; APPLICANT: ADAMS, GREGORY P.
; APPLICANT: HORAK, EVA M.
; APPLICANT: WEINER, LOUIS M.
; APPLICANT: JAMES, MARKS D.
; TITLE OF INVENTION: BISPECIFIC SINGLE CHAIN Fv ANTIBODY MOLECULES AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 407T-000420US
; CURRENT APPLICATION NUMBER: US/11/154,103
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: US 60/370,276
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US10/406,830
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic antibody.
US-11-154-103-10

Query Match 11.9%; Score 31; DB 7; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 GTLVTVSSGGSGGGGGGGSEIVLTQSP 141
Db 138 GTLVTVSSGGSGGGGGGGSEIVLTQSP 168

RESULT 10

US-10-497-088-21
; Sequence 21, Application US/10497088
; Publication No. US20060088520A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: Germersaad, Wilfred
; APPLICANT: Logtenberg, Ton
; APPLICANT: Lekkerkerker, Annemarie N
; TITLE OF INVENTION: Antigen presenting cell targeting conjugate, an antigen
; TITLE OF INVENTION: presenting cell contacted with such conjugate, their use for
; TITLE OF INVENTION: vaccination or as medicament, and methods for their production
; TITLE OF INVENTION: generation
; FILE REFERENCE: 0070 US 00 CON
; CURRENT APPLICATION NUMBER: US/10/497,088
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: PCT/EP01/14255
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/EP02/13681
; PRIOR FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: EP01204997.9
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1052
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MatDC16-Cgamma4-MAGE-1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (546)..(546)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid

US-11-249-296-6
 ; Sequence 6, Application US/11249296
 ; Publication No. US20060115428A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schering Aktiengesellschaft
 ; TITLE OF INVENTION: Identification and Characterization of Function-Blocking
 ; FILE OF INVENTION: Anti-ED-B-Fibronectin Antibodies
 ; FILE REFERENCES: 33042P DE (WWHC)
 ; CURRENT APPLICATION NUMBER: US/11/249,296
 ; CURRENT FILING DATE: 2005-10-14
 ; NUMBER OF SEQ ID NOS: 90
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 117
 ; TYPE: PRT
 ; ORGANISM: human

US-10-570-220-49
 ; Sequence 49, Application US/10570220
 ; Publication No. US20060171941A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The Government of the United States of America, as
 ; APPLICANT: represented by the Secretary of the Department of
 ; APPLICANT: Human Services
 ; APPLICANT: Kashmiri, Syed V.S.
 ; APPLICANT: Schlom, Jeffrey
 ; APPLICANT: Padlan, Eduardo A.
 ; TITLE OF INVENTION: MINIMALLY IMMUNOGENIC VARIANTS OF SDR-GRAN
 ; TITLE OF INVENTION: CC49 AND THEIR USE
 ; FILE REFERENCE: 4239-66649-03
 ; CURRENT APPLICATION NUMBER: US/10570,220
 ; CURRENT FILING DATE: 2006-02-28
 ; PRIOR APPLICATION NUMBER: 60/498,903
 ; PRIOR FILING DATE: 2003-08-29
 ; PRIOR APPLICATION NUMBER: PCT/US04/028004
 ; PRIOR FILING DATE: 2004-08-27
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 49
 ; LENGTH: 30
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: heavy chain framework region 1 of humaniz
 ; US-10-570-220-49

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:40:20 ; Search time 605 Seconds
(without alignments)
658.736 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 8366291 seqs, 1526956180 residues

Word size : 1

Total number of hits satisfying chosen parameters: 8362699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents AA Main.*

1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCTUS_COMB.pep.*
2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066_COMB.pep.*
3: /EMC_Celerra_SIDS3/ptodata/2/paa/US074_COMB.pep.*
4: /EMC_Celerra_SIDS3/ptodata/2/paa/US073_COMB.pep.*
5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075_COMB.pep.*
6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076_COMB.pep.*
7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077_COMB.pep.*
8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078_COMB.pep.*
9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079_COMB.pep.*
10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080_COMB.pep.*
11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081_COMB.pep.*
12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082_COMB.pep.*
13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083_COMB.pep.*
14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084_COMB.pep.*
15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085_COMB.pep.*
16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086_COMB.pep.*
17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087_COMB.pep.*
18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088_COMB.pep.*
19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089_COMB.pep.*
20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090_COMB.pep.*
21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091_COMB.pep.*
22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092_COMB.pep.*
23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093_COMB.pep.*
24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094_COMB.pep.*
25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095_COMB.pep.*
26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096_COMB.pep.*
27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097_COMB.pep.*
28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098_COMB.pep.*
29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099_COMB.pep.*
30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100_COMB.pep.*
31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101_COMB.pep.*
32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102_COMB.pep.*
33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103_COMB.pep.*
34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104_COMB.pep.*
35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105_COMB.pep.*
36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106_COMB.pep.*
37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107_COMB.pep.*
38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108_COMB.pep.*
39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109_COMB.pep.*
40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110_COMB.pep.*
41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111_COMB.pep.*
42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112_COMB.pep.*
43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113_COMB.pep.*

44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114_COMB.pep.*
45: /EMC_Celerra_SIDS3/ptodata/2/paa/US500_COMB.pep.*
46: /EMC_Celerra_SIDS3/ptodata/2/paa/US501_COMB.pep.*
47: /EMC_Celerra_SIDS3/ptodata/2/paa/US502_COMB.pep.*
48: /EMC_Celerra_SIDS3/ptodata/2/paa/US503_COMB.pep.*
49: /EMC_Celerra_SIDS3/ptodata/2/paa/US504_COMB.pep.*
50: /EMC_Celerra_SIDS3/ptodata/2/paa/US505_COMB.pep.*
51: /EMC_Celerra_SIDS3/ptodata/2/paa/US506_COMB.pep.*
52: /EMC_Celerra_SIDS3/ptodata/2/paa/US507_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	261	100.0	261	1	PCT-US05-04612-2	Sequence 2, Appli
2	261	100.0	261	40	US-11-056-825-2	Sequence 2, Appli
3	217	83.1	259	1	PCT-US05-04612-7	Sequence 7, Appli
4	217	83.1	259	40	US-11-056-825-7	Sequence 7, Appli
5	48	18.4	100	1	PCT-US02-22011-75	Sequence 75, Appl
6	48	18.4	100	1	PCT-US02-38450-22	Sequence 22, Appl
7	48	18.4	100	1	PCT-US03-06598-87	Sequence 87, Appl
8	48	18.4	100	31	US-10-153-382-36	Sequence 36, Appl
9	48	18.4	100	33	US-10-308-817-22	Sequence 22, Appl
10	48	18.4	100	33	US-10-379-392-87	Sequence 87, Appl
11	48	18.4	100	34	US-10-453-698-22	Sequence 22, Appl
12	48	18.4	100	36	US-10-612-497-113	Sequence 113, App
13	48	18.4	100	36	US-10-612-497A-113	Sequence 113, App
14	48	18.4	100	36	US-10-683-733-45	Sequence 45, Appl
15	48	18.4	100	37	US-10-776-649-113	Sequence 113, App
16	48	18.4	100	40	US-11-004-590-82	Sequence 82, Appl
17	48	18.4	100	40	US-11-054-669-75	Sequence 75, Appl
18	48	18.4	100	40	US-11-084-554-107	Sequence 107, App
19	48	18.4	100	40	US-11-085-368-36	Sequence 36, Appl
20	48	18.4	100	41	US-11-128-900-113	Sequence 113, App
21	48	18.4	100	41	US-11-136-250-107	Sequence 107, App
22	48	18.4	100	42	US-11-239-308-45	Sequence 45, Appl
23	48	18.4	100	42	US-11-239-308A-45	Sequence 45, Appl
24	48	18.4	104	27	US-09-791-537-107428	Sequence 107428, A
25	48	18.4	108	27	US-09-791-537-23019	Sequence 23019, A
26	48	18.4	110	27	US-09-791-537-24305	Sequence 24305, A
27	48	18.4	110	46	US-60-164-762-847	Sequence 847, App
28	48	18.4	111	1	PCT-US01-03537-11	Sequence 11, Appl
29	48	18.4	111	1	PCT-US01-03537-59	Sequence 59, Appl
30	48	18.4	111	1	PCT-US01-12139-2	Sequence 2, Appli
31	48	18.4	111	1	PCT-US03-39599A-11	Sequence 11, Appl
32	48	18.4	111	1	PCT-US03-39599A-59	Sequence 59, Appl
33	48	18.4	111	28	US-09-835-087-2	Sequence 2, Appli
34	48	18.4	111	36	US-10-662-061-13	Sequence 13, Appl
35	48	18.4	111	37	US-10-706-852-15	Sequence 15, Appl
36	48	18.4	111	37	US-10-733-563-11	Sequence 11, Appl
37	48	18.4	111	37	US-10-733-563-59	Sequence 59, Appl
38	48	18.4	111	37	US-10-766-610-11	Sequence 11, Appl
39	48	18.4	111	37	US-10-766-610-59	Sequence 59, Appl
40	48	18.4	111	37	US-10-766-773-11	Sequence 11, Appl
41	48	18.4	111	37	US-10-766-773-59	Sequence 59, Appl
42	48	18.4	112	1	PCT-US01-03537-56	Sequence 56, Appl
43	48	18.4	112	1	PCT-US01-03537-62	Sequence 62, Appl
44	48	18.4	112	1	PCT-US01-03537-66	Sequence 66, Appl
45	48	18.4	112	1	PCT-US03-39599A-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1

PCT-US05-04612-2

; Sequence 2, Application PC/TUS0504612

; GENERAL INFORMATION:

```

; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Sayen, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0043
; CURRENT APPLICATION NUMBER: PCT/US05/04612
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: US 11/056,825
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
PCT-US05-04612-2

Query Match      100.0%; Score 261; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.2e-236;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
QY 181 SPRRLIYKVSNRDGVDPDRFSGSGGTDTLTKISRVEADIGVYCMQGTTHWPPRTFGQ 240
DB 181 SPRRLIYKVSNRDGVDPDRFSGSGGTDTLTKISRVEADIGVYCMQGTTHWPPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDDK 261

RESULT 2
US-11-056-825-2
; Sequence 2, Application US/11056825
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Sayen, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-2

Query Match      100.0%; Score 261; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.2e-236;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
QY 181 SPRRLIYKVSNRDGVDPDRFSGSGGTDTLTKISRVEADIGVYCMQGTTHWPPRTFGQ 240
DB 181 SPRRLIYKVSNRDGVDPDRFSGSGGTDTLTKISRVEADIGVYCMQGTTHWPPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDDK 261

RESULT 2
US-11-056-825-2
; Sequence 2, Application US/11056825
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Sayen, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-2
```

```

Query Match      100.0%; Score 261; DB 40; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.2e-236;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
QY 181 SPRRLIYKVSNRDGVDPDRFSGSGGTDTLTKISRVEADIGVYCMQGTTHWPPRTFGQ 240
DB 181 SPRRLIYKVSNRDGVDPDRFSGSGGTDTLTKISRVEADIGVYCMQGTTHWPPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDDK 261

RESULT 3
PCT-US05-04612-7
; Sequence 7, Application PC/TUS0504612
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Sayen, Alan
; FILE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0043
; CURRENT APPLICATION NUMBER: PCT/US05/04612
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: US 11/056,825
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
PCT-US05-04612-7

Query Match      83.1%; Score 217; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.9e-195;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKKPGASVKVCKASGYTFSNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFOQRPGQ 180
QY 181 SPRRLIYKVSNRDGVDPDRFSGSGGTDTLTKISRVE 217
DB 181 SPRRLIYKVSNRDGVDPDRFSGSGGTDTLTKISRVE 217
```

```
RESULT 4
US-11-056-825-7
; Sequence 7, Application US/11056825
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-7

Query Match      83.1%; Score 217; DB 40; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.9e-195;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKSCASGYTFSTNYGITWVRQAPGGLEWMGWINNNTYH 60
Db 1 MAQVQLVQSGAEVKKPGASVKSCASGYTFSTNYGITWVRQAPGGLEWMGWINNNTYH 60

Qy 61 AQKQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 AQKQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120

Qy 121 GSGGGGGGGGGSEIVLTQSLPLPVLTLGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180
Db 121 GSGGGGGGGGGSEIVLTQSLPLPVLTLGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180

Qy 181 SPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVE 217
Db 181 SPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVE 217

RESULT 5
PCT-US02-22011-75
; Sequence 75, Application PC/TUS0222011
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.02
; CURRENT APPLICATION NUMBER: PCT/US02/22011
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-22011-75

Query Match      18.4%; Score 48; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 220
Db 40 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 87

RESULT 6
PCT-US02-38450-22
; Sequence 22, Application PC/TUS0238450
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: PCT/US02/38450
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
PCT-US02-38450-22

Query Match      18.4%; Score 48; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 220
Db 40 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 87

RESULT 7
PCT-US03-06598-87
; Sequence 87, Application PC/TUS0306598
; GENERAL INFORMATION:
; APPLICANT: XENCOR
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: FP-71386-3 463077-237
; CURRENT APPLICATION NUMBER: PCT/US03/06598
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-06598-87

Query Match      18.4%; Score 48; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 220
Db 40 WFQORPGQSPRLIYKVNRRDGVDPDRFSGSGSGTDFTLKISRVEAED 87

RESULT 8
US-10-153-382-36
; Sequence 36, Application US/10153382
; GENERAL INFORMATION:
; APPLICANT: PFIZER PRODUCTS INC.
; TITLE OF INVENTION: USES OF ANTI-CTLA-4 ANTIBODIES
; FILE REFERENCE: PC23019A
; CURRENT APPLICATION NUMBER: US/10/153,382
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: 60/293042
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 39
```



```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-382-36

Query Match      18.4%; Score 48; DB 31; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 220
Db 40 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 87

RESULT 9
US-10-308-817-22
; Sequence 22, Application US/10308817
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; APPLICANT: Wu, Davang
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
US-10-308-817-22

Query Match      18.4%; Score 48; DB 33; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 220
Db 40 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 87

RESULT 10
US-10-379-392-87
; Sequence 87, Application US/10379392
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiya, Bassil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: A-71386-3 463077-236
; CURRENT APPLICATION NUMBER: US/10/379,392
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 87
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-379-392-87

Query Match      18.4%; Score 48; DB 33; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 220
Db 40 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 87
```

```
Db 40 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 87

RESULT 11
US-10-453-698-22
; Sequence 22, Application US/10453698
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 82 CIP (1087-37 CIP)
; CURRENT APPLICATION NUMBER: US/10/453,698
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 100
; TYPE: PRT
; ORGANISM: human
US-10-453-698-22

Query Match      18.4%; Score 48; DB 34; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 220
Db 40 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 87

RESULT 12
US-10-612-497-113
; Sequence 113, Application US/10612497
; GENERAL INFORMATION:
; APPLICANT: Douglas C. Hanson
; APPLICANT: Mark J. Neveu
; APPLICANT: Eileen E. Mueller
; APPLICANT: Jeffrey H. Hanke
; APPLICANT: Steven C. Gilman
; APPLICANT: C. Geoffrey Davis
; APPLICANT: Jose R. Corvalan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PFI DIVI
; CURRENT APPLICATION NUMBER: US/10/612,497
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-497-113

Query Match      18.4%; Score 48; DB 36; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 220
Db 40 WFOQRPQGSPPRLIYKVSNRDSGVDPDRFSGSGGTDTTLKISRVEAD 87

RESULT 13
US-10-612-497A-113
; Sequence 113, Application US/10612497A
; GENERAL INFORMATION:
; APPLICANT: HANSON, DOUGLAS C.
; APPLICANT: NEVEU, MARK J.
; APPLICANT: MUELLER, EILEEN E.
; APPLICANT: HANKE, JEFFREY H.
```

; APPLICANT: GILMAN, STEVEN C.
; APPLICANT: DAVIS, C. GEOFFREY
; APPLICANT: CORVALAN, JOSE R.
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PF1
; CURRENT APPLICATION NUMBER: US/10/612,497A
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US/09/472,087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-612-497A-113

Query Match 18.4%; Score 48; DB 36; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 87

RESULT 14

US-10-683-733-45
; Sequence 45, Application US/10683733
; GENERAL INFORMATION:
; APPLICANT: Smider, Vaughn
; APPLICANT: Larrick, James W.
; APPLICANT: Integrigen, Inc.
; TITLE OF INVENTION: Recombinant Catalytic Polypeptides and Their Uses
; FILE REFERENCE: 021216-000310US
; CURRENT APPLICATION NUMBER: US/10/683,733
; CURRENT FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: US 60/417,979
; PRIOR FILING DATE: 2002-10-09
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-683-733-45

Query Match 18.4%; Score 48; DB 36; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 87

RESULT 15

US-10-776-649-113
; Sequence 113, Application US/10776649
; GENERAL INFORMATION:
; APPLICANT: Douglas C. Hanson
; APPLICANT: Mark J. Neveu
; APPLICANT: Eileen B. Mueller
; APPLICANT: Jeffrey H. Hanke
; APPLICANT: Steven C. Gilman
; APPLICANT: C. Geoffrey Davis
; APPLICANT: Jose R. Corvalan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PF1 Div2
; CURRENT APPLICATION NUMBER: US/10/776,649
; CURRENT FILING DATE: 2004-02-10

; PRIOR APPLICATION NUMBER: US 10/612497
; PRIOR FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: US 09/472087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-776-649-113

Query Match 18.4%; Score 48; DB 37; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 220
Db 40 WFOQPGQSPRRLIYKVSNRDSGVDPDRFSGSGCTDFTLKISRVEAED 87

Search completed: August 9, 2006, 13:51:13
Job time : 607 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:41:18 ; Search time 35 Seconds
(without alignments)
496.727 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVYKDDDDK 261

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 294602 seqs, 66610880 residues

Word size : 1

Total number of hits satisfying chosen parameters: 293948

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/paa/US11_NEW_COMB.pep.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	18.4	100	7	US-11-432-734-75
2	48	18.4	100	7	US-11-332-619-87
3	48	18.4	103	7	US-11-346-403A-38
4	48	18.4	111	6	US-10-377-122-12
5	48	18.4	253	8	US-60-797-351-1964
6	48	18.4	253	8	US-60-814-869-1964
7	48	18.4	253	8	US-60-815-559-1964
8	48	18.4	253	8	US-60-834-152-1964
9	39	14.9	246	8	US-60-797-351-1192
10	39	14.9	246	8	US-60-814-869-1192
11	39	14.9	246	8	US-60-815-559-1192
12	39	14.9	246	8	US-60-834-152-1192
13	37	14.2	247	8	US-60-797-351-3242
14	37	14.2	247	8	US-60-814-869-3242
15	37	14.2	247	8	US-60-815-559-3242
16	37	14.2	247	8	US-60-834-152-3242
17	36	13.8	122	7	US-11-411-003-374
18	36	13.8	141	7	US-11-411-003-305
19	36	13.8	255	8	US-60-797-351-1190
20	36	13.8	255	8	US-60-814-869-1190
21	36	13.8	255	8	US-60-815-559-1190
22	36	13.8	255	8	US-60-834-152-1190
23	35	13.4	117	1	PCT-US06-16786-118
24	35	13.4	117	7	US-11-433-924-118
25	35	13.4	118	1	PCT-US06-16786-198

26	35	13.4	118	7	US-11-433-924-198	Sequence 198, App
27	35	13.4	119	1	PCT-US06-16786-406	Sequence 406, App
28	35	13.4	119	7	US-11-411-003-338	Sequence 338, App
29	35	13.4	119	7	US-11-411-003-370	Sequence 370, App
30	35	13.4	119	7	US-11-433-924-406	Sequence 406, App
31	35	13.4	126	1	PCT-US06-05691-4	Sequence 4, Appli
32	35	13.4	138	7	US-11-411-003-321	Sequence 321, App
33	35	13.4	251	8	US-60-797-351-1315	Sequence 1315, Ap
34	35	13.4	251	8	US-60-814-869-1315	Sequence 1315, Ap
35	35	13.4	251	8	US-60-815-559-1315	Sequence 1315, Ap
36	35	13.4	251	8	US-60-834-152-1315	Sequence 1315, Ap
37	35	13.4	253	8	US-60-797-351-1611	Sequence 1611, Ap
38	35	13.4	253	8	US-60-814-869-1611	Sequence 1611, Ap
39	35	13.4	253	8	US-60-815-559-1611	Sequence 1611, Ap
40	35	13.4	254	8	US-60-797-351-1779	Sequence 1779, Ap
41	35	13.4	254	8	US-60-814-869-1779	Sequence 1779, Ap
42	35	13.4	254	8	US-60-815-559-1779	Sequence 1779, Ap
43	35	13.4	254	8	US-60-814-869-1780	Sequence 1780, Ap
44	35	13.4	254	8	US-60-815-559-1780	Sequence 1780, Ap
45	35	13.4	254	8	US-60-815-559-1779	Sequence 1779, Ap

ALIGNMENTS

RESULT 1
US-11-432-734-75
; Sequence 75, Application US/11432734
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: SUPER HUMANIZED ANTIBODIES
; FILE REFERENCE: 30219/US/3
; CURRENT APPLICATION NUMBER: US/11/432,734
; CURRENT FILING DATE: 2006-05-10
; PRIOR APPLICATION NUMBER: US/11/054,669
; PRIOR FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 10/194,975
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 75
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-432-734-75

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.4e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 WFOBPQSGSPRLIYKVNRSVDPRFSGSGSGDTFTLKISRVEAED 220
DB 40 WFOBPQSGSPRLIYKVNRSVDPRFSGSGSGDTFTLKISRVEAED 87

RESULT 2
US-11-332-619-87
; Sequence 87, Application US/11332619
; GENERAL INFORMATION:
; APPLICANT: Moore, Gregory L.
; TITLE OF INVENTION: ANTIBODIES AND FC FUSION PROTEINS WITH ALTERED IMMUNOGENICITY
; FILE REFERENCE: 187350/US/4 463077-427
; CURRENT APPLICATION NUMBER: US/11/332,619
; CURRENT FILING DATE: 2006-01-12
; PRIOR APPLICATION NUMBER: US 60/643,313
; PRIOR FILING DATE: 2005-01-12
; PRIOR APPLICATION NUMBER: US 60/652,958
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: US 60/654,636
; PRIOR FILING DATE: 2005-02-17

; NUMBER OF SEQ ID NOS: 1939
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 87
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-332-619-87

Query Match 18.4%; Score 48; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.4e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 3
US-11-346-403A-38
; Sequence 38, Application US/11346403A
; GENERAL INFORMATION:
; APPLICANT: Hoet, Rene
; APPLICANT: Schoonbroodt, Sonia
; APPLICANT: Ladner, Robert C.
; TITLE OF INVENTION: LIBRARIES AND METHODS FOR
; TITLE OF INVENTION: ISOLATING ANTIBODIES
; FILE REFERENCE: 10280-116001
; CURRENT APPLICATION NUMBER: US/11/346.403A
; CURRENT FILING DATE: 2006-02-01
; PRIOR APPLICATION NUMBER: US 60/649,065
; PRIOR FILING DATE: 2005-02-01
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-11-346-403A-38

Query Match 18.4%; Score 48; DB 7; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 41 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 88

RESULT 4
US-10-377-122-12
; Sequence 12, Application US/10377122
; GENERAL INFORMATION:
; APPLICANT: HANSEN, HANS J.
; APPLICANT: LEUNG, SHUI-ON
; APPLICANT: OU, ZHENGXING
; APPLICANT: GOLDENBERG, DAVID M.
; TITLE OF INVENTION: INTERNALIZING ANTI-CD-74 ANTIBODIES AND METHODS OF USE
; FILE REFERENCE: 78258/329656
; CURRENT APPLICATION NUMBER: US/10/377,122
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/360,259
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-377-122-12

Query Match 18.4%; Score 48; DB 6; Length 111;

Best Local Similarity 100.0%; Pred. No. 2.6e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 40 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 87

RESULT 5
US-60-797-351-1964
; Sequence 1964, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1964

Query Match 18.4%; Score 48; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 180 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 227

RESULT 6
US-60-814-869-1964
; Sequence 1964, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-1964

Query Match 18.4%; Score 48; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 173 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 220
Db 180 WFOQRPQGSPRRLIYKVSNRDGVDPDRFSGSGGTDFTLKISRVEAED 227

RESULT 7
US-60-815-559-1964
; Sequence 1964, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-1964

```

Query Match      18.4%; Score 48; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDGVDPDRFSGSGTDFTLKISRVEAED 220
Db 180 WFOQRPQSPRLIYKVSNRDGVDPDRFSGSGTDFTLKISRVEAED 227

RESULT 8
US-60-834-152-1964
; Sequence 1964, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1964

Query Match      18.4%; Score 48; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQRPQSPRLIYKVSNRDGVDPDRFSGSGTDFTLKISRVEAED 220
Db 180 WFOQRPQSPRLIYKVSNRDGVDPDRFSGSGTDFTLKISRVEAED 227

RESULT 9
US-60-797-351-1192
; Sequence 1192, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1192
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1192

Query Match      14.9%; Score 39; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 98
Db 59 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 97

RESULT 10
US-60-814-869-1192
; Sequence 1192, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1192
; LENGTH: 246
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-60-814-869-1192

Query Match      14.9%; Score 39; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 98
Db 59 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 97

RESULT 11
US-60-815-559-1192
; Sequence 1192, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1192
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-1192

Query Match      14.9%; Score 39; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 98
Db 59 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 97

RESULT 12
US-60-834-152-1192
; Sequence 1192, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1192
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1192

Query Match      14.9%; Score 39; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 98
Db 59 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCAR 97

RESULT 13
US-60-797-351-3242
; Sequence 3242, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 3242

```

```
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-3242

Query Match      14.2%; Score 37; DB 8; Length 247;
Best Local Similarity 100.0%; Pred. No. 9.7e-27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|
Db      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|

RESULT 14
US-60-814-869-3242
; Sequence 3242, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 3242
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-3242

Query Match      14.2%; Score 37; DB 8; Length 247;
Best Local Similarity 100.0%; Pred. No. 9.7e-27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|
Db      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|

RESULT 15
US-60-815-559-3242
; Sequence 3242, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 3242
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-3242

Query Match      14.2%; Score 37; DB 8; Length 247;
Best Local Similarity 100.0%; Pred. No. 9.7e-27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|
Db      62 QKFGQGVMTTDTSTSTAYMELSLRSLRSDDTAVYYCAR 98
|

Search completed: August 9, 2006, 13:51:53
Job time : 36 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:34:32 ; Search time 39 Seconds
(without alignments)
643.912 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 261

Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGGLVDYKDDDK 261

Scoring table:

OLIGO Gap 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 1

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	18.4	114	2 S49572	Ig kappa chain pre
2	48	18.4	114	2 B49002	Ig kappa chain V r
3	48	18.4	120	2 S42268	Ig kappa chain V r
4	48	18.4	133	1 A24452	Ig kappa chain pre
5	48	18.4	133	1 K2HURP	Ig kappa chain pre
6	48	18.4	133	2 S23230	Ig kappa chain pre
7	48	18.4	133	2 S42611	HUNVK protein prec
8	48	18.4	142	2 S22902	Ig kappa chain V r
9	37	14.2	122	2 S40338	Ig kappa chain - h
10	35	13.4	129	2 S36260	Ig heavy chain V r
11	34	13.0	98	2 S26919	Ig heavy chain V r
12	34	13.0	131	2 S21924	Ig heavy chain V r
13	33	12.6	122	2 S36271	Ig heavy chain V r
14	33	12.6	132	2 S40322	Ig kappa chain - h
15	32	12.3	118	2 S40374	Ig kappa chain - h
16	32	12.3	140	2 S19655	Ig heavy chain pre
17	32	12.3	140	2 S22658	Ig heavy chain V r
18	29	11.1	98	2 S26920	Ig heavy chain V r
19	29	11.1	98	2 S26918	Ig heavy chain V r
20	29	11.1	98	2 S26912	Ig heavy chain V r
21	29	11.1	98	2 S26938	Ig heavy chain V r
22	29	11.1	98	2 S26913	Ig heavy chain V r
23	29	11.1	104	2 S69859	Ig heavy chain V r
24	29	11.1	111	2 S21925	Ig heavy chain V r
25	29	11.1	116	2 S31667	Ig heavy chain V r
26	29	11.1	117	1 HVH035	Ig heavy chain pre
27	29	11.1	117	1 HVH035	Ig heavy chain pre
28	29	11.1	117	2 S18552	Ig heavy chain V r
29	29	11.1	117	2 S31680	Ig heavy chain V r

30	29	11.1	117	2 S18553	Ig heavy chain V r
31	29	11.1	117	2 S18551	Ig heavy chain V r
32	29	11.1	118	2 S36265	Ig heavy chain V r
33	29	11.1	129	2 S43393	Ig heavy chain V r
34	29	11.1	132	2 S31596	Ig heavy chain V r
35	29	11.1	135	2 S49530	anti-Sm antibody V
36	29	11.1	142	2 A32483	Ig heavy chain V r
37	29	11.1	148	2 S29257	Ig heavy chain V r
38	28	10.7	120	2 S42267	Ig kappa chain V r
39	27	10.3	54	2 S34093	Ig kappa chain V r
40	27	10.3	71	2 H30538	Ig kappa chain V-J
41	27	10.3	75	2 S40337	Ig kappa chain V-J
42	27	10.3	87	2 S34094	Ig kappa chain V r
43	27	10.3	87	2 S34091	Ig kappa chain V r
44	27	10.3	93	2 PH1039	Ig light chain V r
45	27	10.3	94	2 PL0258	Ig kappa chain V r

ALIGNMENTS

RESULT 1

S49572 Ig kappa chain precursor - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 06-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 21-Jan-2000

C:Accession: S49572

R:Giachino, C.; Padova, E.; Lanzavecchia, A.

A:Description: k+l+ dual receptor B cells are present in the human peripheral repertoire.

A:Reference number: S49571

A:Accession: S49572

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-114 <GIA>

A:Cross-references: UNIPARC:UPI0000116709; EMBL:Z46626; NID:G575261; PIDN:CAR86596.1; PII

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F16-95/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQRPQSGSPRLIYKVSNRDSGVDPFRFSGSGSGTDTLKISRVEAED 220

Db 40 WFQRPQSGSPRLIYKVSNRDSGVDPFRFSGSGSGTDTLKISRVEAED 87

RESULT 2

B49002

Ig kappa chain V region, rheumatoid factor RF antibody - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: B49002

R:Stuber, F.; Lee, S.K.; Bridges Jr., S.L.; Koopman, W.J.; Schroeder, H.W.J.; Gaskin, F.;

Arthritis Rheum. 35, 900-904, 1992

A:Title: A rheumatoid factor from a normal individual encoded by VH2 and V kappa II gene

A:Reference number: A49002; MUID:92352481; PMID:1322670

A:Accession: B49002

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-114 <STU>

A:Cross-references: UNIPROT:Q9UL80; UNIPARC:UPI0000176B43

A:Experimental source: EBV-transformed lymphoblastoid cell line SSH23

A>Note: sequence inconsistent with the nucleotide translation

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F16-95/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 40 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 3
S42268
Ig kappa chain V region (A17) - human
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 31-Dec-2004
C:Accession: S42268
R:Lautner-Rieske, A.; Huber, C.; Meindl, A.; Pargent, W.; Schaeuble, K.F.; Thiebe, R.; Zolner, J. Immunol. 22, 1023-1029, 1992
A:Title: The human immunoglobulin kappa locus. Characterization of the duplicated A region
A:Reference number: A49043; MUID:92201291; PMID:1551402
A:Accession: S42268
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <LAU>
A:Cross-references: UNIPROT:Q8TCD0; UNIPARC:UPI0000176EB9; EMBL:X63403
C:Genetics:
A:Introns: 17/1
C:Superfamily: immunoglobulin homology
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 4
A24452
Ig kappa chain precursor V-II region (RPMI 6410) - human
C:Species: Homo sapiens (man)
C>Date: 24-Jan-1988 #sequence_revision 09-Aug-1996 #text_change 16-Jul-1999
C:Accession: A24452
R:Weir, L.; Leder, P.
Nucleic Acids Res. 14, 3957-3970, 1986
A:Title: Structure and expression of a human subgroup II immunoglobulin kappa gene.
A:Reference number: A24452; MUID:86232631; PMID:3086847
A:Accession: A24452
A:Molecule type: DNA
A:Residues: 1-133 <WEI>
A:Cross-references: UNIPARC:UPI0000113B46; GB:M36859; NID:g185932; PIDN:AAA58920.1; PID:
A:Note: this sequence was determined from the differentiated gene
C:Genetics:
A:Gene: GDB:IGKV2
A:Cross-references: GDB:136265
A:Map position: 2p12-2p12
A:Introns: 17/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:21-20/Domain: signal sequence #status predicted <SIG>
F:21-133/Product: Ig kappa chain V-II region (RPMI 6410) #status predicted <MAT>
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 5
K2HURP
Ig kappa chain precursor V-II region (RPMI) - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A01890
R:Klobeck, H.G.; Meindl, A.; Combrinato, G.; Solomon, A.; Zachau, H.G.
Nucleic Acids Res. 13, 6499-6513, 1985
A:Title: Human immunoglobulin kappa light chain genes of subgroups II and III.
A:Reference number: A93588; MUID:86041852; PMID:2997711
A:Accession: A01890
A:Molecule type: DNA
A:Residues: 1-133 <KLO>
A:Cross-references: UNIPROT:P06310; UNIPARC:UPI000012E159
A:Note: the sequence was determined from the differentiated gene
C:Genetics:
A:Gene: GDB:IGKV2
A:Cross-references: GDB:136265
A:Map position: 2p12-2p12
A:Introns: 17/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:21-20/Domain: signal sequence #status predicted <SIG>
F:21-133/Product: Ig kappa chain V-II region (RPMI) #status predicted <MAT>
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 6
S23230
Ig kappa chain precursor V-J region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S23230
R:Kennedy, M.A.
J. Exp. Med. 173, 1033-1036, 1991
A:Title: Novel chromosome translocation caused by fusion of immunoglobulin heavy and light chain genes.
A:Reference number: S23230; MUID:91178438; PMID:1840606
A:Accession: S23230
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <KEN>
A:Cross-references: UNIPARC:UPI0000115EA9; EMBL:X55400; NID:g33999; PIDN:CAA39072.1; PID:
C:Genetics:
A:Introns: 17/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:36-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFQORPGQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAED 107

```

RESULT 7
S42611
HUNK protein precursor - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S42611
R;Spatz, L.A.; Williams, M.; Brender, B.; Desai, R.; Latov, N.
J. Neuroimmunol. 36, 29-39, 1992
A:Title: DNA sequence analysis and comparison of the variable heavy and light chain regions
A:Reference number: S42610; MUID:92138794; PMID:1370957
A:Accession: S42611
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <SPA>
A:Cross-references: UNIPARC:UPI000011378B; EMBL:X54137; NID:G433889; PIDN:CAA38072.1; PID:
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F;35-115/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.9e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRRLLYKVSNRDSGVDPFRSGSGSGTDTFLKISRVEAED 220
Db 60 WFOQPGQSPRRLLYKVSNRDSGVDPFRSGSGSGTDTFLKISRVEAED 107

RESULT 8
S22902
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S22902
R;Chastagner, P.; Theze, J.; Zouali, M.
Gene 101, 305-306, 1991
A:Title: Cloning of a gene encoding a lupus-associated human autoantibody V(K) region us
A:Reference number: S22902; MUID:91276289; PMID:1905262
A:Accession: S22902
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-142 <CHA>
A:Cross-references: UNIPARC:UPI0000176CAB; EMBL:X56510
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F;47-126/Domain: immunoglobulin homology <IMM>

Query Match 18.4%; Score 48; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.9e-36;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRRLLYKVSNRDSGVDPFRSGSGSGTDTFLKISRVEAED 220
Db 71 WFOQPGQSPRRLLYKVSNRDSGVDPFRSGSGSGTDTFLKISRVEAED 118

RESULT 9
S40338
Ig kappa chain - human
C:Species: Homo sapiens (man)
C>Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S40338
R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A:Title: Expressed human immunoglobulin chi genes and their hypermutation.
A:Reference number: S40312; MUID:94080891; PMID:8258341
A:Accession: S40338
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-122 <KLE>
A:Cross-references: UNIPARC:UPI0000116158; EMBL:X72448; NID:G441364; PIDN:CAA51116.1; PID:
C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin
F;22-101/Domain: immunoglobulin homology <IMM>

Query Match 14.2%; Score 37; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 2e-26;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 135 IVLTQSPISLPVTLGPASISCRSSQNLVYSDGNTYL 171
Db 8 IVLTQSPISLPVTLGPASISCRSSQNLVYSDGNTYL 44

RESULT 10

S36260
Ig heavy chain V region (clone alpha-CEA4-8A) - human (fragment)
C:Species: Homo sapiens (man)

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999

C:Accession: S36260

R;Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.;
EMBO J. 12, 725-734, 1993

A:Title: Human anti-self antibodies with high specificity from phage display libraries.

A:Reference number: S36256; MUID:93178448; PMID:7679990

A:Accession: S36260

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-129 <GRI>

A:Cross-references: UNIPARC:UPI0000118DEB; EMBL:Z18851; NID:G33124; PIDN:CAA79303.1; PID:
C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 13.4%; Score 35; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.4e-24;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 QQRVTMTTDTSTAYMELSLRSLRSDDTAVYYCARD 99

Db 65 QQRVTMTTDTSTAYMELSLRSLRSDDTAVYYCARD 99

RESULT 11

S26919

Ig heavy chain V region (DP-14) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S26919

R;Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.

J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V.

A:Reference number: S26885; MUID:93021117; PMID:1404388

A:Accession: S26919

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-98 <ROM>

A:Cross-references: UNIPARC:UPI0000031F31; EMBL:Z12316; NID:G32855; PIDN:CAA78186.1; PID:

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 13.0%; Score 34; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 QQRVTMTTDTSTAYMELSLRSLRSDDTAVYYCAR 98

Db 65 QQRVTMTTDTSTAYMELSLRSLRSDDTAVYYCAR 98

RESULT 12

S21924

Ig heavy chain V region - human

C:Species: Homo sapiens (man)

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999

C;Accession: S21924; S21923
R;Friedman, D.F.
submitted to the EMBL Data Library, July 1991
A;Reference number: S21923
A;Accession: S21924
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-131 <FRI>
A;Cross-references: UNIPARC:UPI0000115FA2; EMBL:X60505; NID:g33565; PIDN:CAA43025.1; PID
C;Genetics:
A;Introns: 16/1
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;34-117/Domain: immunoglobulin homology <IMM>

Query Match 13.0%; Score 34; DB 2; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 QGRVTMTTDTSTAYMELRLSRSDDTAVYYCAR 98
|||||
Db 84 QGRVTMTTDTSTAYMELRLSRSDDTAVYYCAR 117
|||||

RESULT 13
S36271
Ig heavy chain V region (clone alpha-THY-29) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C;Accession: S36271
R;Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.
EMBO J. 12, 725-734, 1993
A;Title: Human anti-self antibodies with high specificity from phage display libraries.
A;Reference number: S36256; MUID:93178448; PMID:7679990
A;Accession: S36271
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-122 <GRI>
A;Cross-references: UNIPARC:UPI0000118DE3; EMBL:Z18832; NID:g33115; PIDN:CAA79284.1; PID
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 12.6%; Score 33; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 9.1e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 QGRVTMTTDTSTAYMELRLSRSDDTAVYYCA 97
|||||
Db 65 QGRVTMTTDTSTAYMELRLSRSDDTAVYYCA 97
|||||

RESULT 14
S40322
Ig kappa chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C;Accession: S40322
R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A;Title: Expressed human immunoglobulin chi genes and their hypermutation.
A;Reference number: S40312; MUID:94080891; PMID:8258341
A;Accession: S40322
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-132 <KLE>
A;Cross-references: UNIPARC:UPI0000116148; EMBL:X72432; NID:g441332; PIDN:CAA51100.1; PI
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;35-114/Domain: immunoglobulin homology <IMM>

Query Match 12.6%; Score 33; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 9.8e-23;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 188 KVSNRDSGVDPDRFSGSGGTFTLKISRVEAED 220
|||||
Db 74 KVSNRDSGVDPDRFSGSGGTFTLKISRVEAED 106
|||||

RESULT 15
S40374
Ig kappa chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 31-Dec-2004
C;Accession: S40374
R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A;Title: Expressed human immunoglobulin chi genes and their hypermutation.
A;Reference number: S40312; MUID:94080891; PMID:8258341
A;Accession: S40374
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-118 <KLE>
A;Cross-references: UNIPROT:Q8TCD0; UNIPARC:UPI0000176C9D; EMBL:X72484
C;Superfamily: immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;20-99/Domain: immunoglobulin homology <IMM>

Query Match 12.3%; Score 32; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 7.3e-22;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 VSNRDSGVDPDRFSGSGGTFTLKISRVEAED 220
|||||
Db 60 VSNRDSGVDPDRFSGSGGTFTLKISRVEAED 91
|||||

Search completed: August 9, 2006, 13:40:06
Job time : 40 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:31:07 ; Search time 301 Seconds
(without alignments)
802.090 Million cell updat

Title: US-11-056-825-2

Perfect score:

Sequence: 1 MAQVQLVQSGAEVKKPGASV.....KVEIKRGLGGLVDYKDDDDK 261

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size : 1

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt 7.2:*

```
database :
  uniprot : 7.2,
  1: uniprot_sprot:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	48	18.4	86	2	Q723Y5	HUMAN
2	48	18.4	114	2	Q9UL80	HUMAN
3	48	18.4	133	1	KV2F	HUMAN
4	41	15.7	239	2	Q8TCD0	HUMAN
5	33	12.6	500	2	Q6N091	HUMAN
6	32	12.3	159	2	Q96QS0	HUMAN
7	29	11.1	117	1	HV1B	HUMAN
8	29	11.1	117	1	HV1G	HUMAN
9	27	10.3	112	2	Q53VP8	MOUSE
10	27	10.3	113	1	KV2D	HUMAN
11	27	10.3	113	1	KV2G	MOUSE
12	27	10.3	115	2	Q5F2I0	MOUSE
13	27	10.3	117	1	KV2E	HUMAN
14	27	10.3	219	2	Q65ZC0	MOUSE
15	27	10.3	239	2	Q58EU8	MOUSE
16	27	10.3	240	2	Q6PIH6	HUMAN
17	27	10.3	248	2	Q65ZQ7	9MURI
18	26	10.0	244	2	Q65ZC8	HUMAN
19	24	9.2	124	2	Q9UL92	HUMAN
20	24	9.2	150	2	Q9Y298	HUMAN
21	23	8.8	115	1	KV2A	HUMAN
22	23	8.8	119	2	Q9UL94	HUMAN
23	23	8.8	125	2	Q9UL95	HUMAN
24	23	8.8	240	2	Q65ZC9	HUMAN
25	23	8.8	519	2	Q5EBM2	HUMAN
26	21	8.0	170	2	Q925S2	MOUSE
27	21	8.0	239	2	Q9NEK0	HUMAN
28	20	7.7	241	2	Q921A6	MOUSE
29	19	7.3	243	2	Q7TQM2	MOUSE
30	19	7.3	255	2	Q6KB05	MOUSE
31	19	7.3	498	2	Q6N041	HUMAN

Accession	Species	Length	GC	GC3	GC3+4	GC3+5	GC3+6	GC3+7	GC3+8	GC3+9	GC3+10	GC3+11	GC3+12	GC3+13	GC3+14	GC3+15	GC3+16	GC3+17	GC3+18	GC3+19	GC3+20	GC3+21	GC3+22	GC3+23	GC3+24	GC3+25	GC3+26	GC3+27	GC3+28	GC3+29	GC3+30	GC3+31	GC3+32	GC3+33	GC3+34	GC3+35	GC3+36	GC3+37	GC3+38	GC3+39	GC3+40	GC3+41	GC3+42	GC3+43	GC3+44	GC3+45	GC3+46	GC3+47	GC3+48	GC3+49	GC3+50	GC3+51	GC3+52	GC3+53	GC3+54	GC3+55	GC3+56	GC3+57	GC3+58	GC3+59	GC3+60	GC3+61	GC3+62	GC3+63	GC3+64	GC3+65	GC3+66	GC3+67	GC3+68	GC3+69	GC3+70	GC3+71	GC3+72	GC3+73	GC3+74	GC3+75	GC3+76	GC3+77	GC3+78	GC3+79	GC3+80	GC3+81	GC3+82	GC3+83	GC3+84	GC3+85	GC3+86	GC3+87	GC3+88	GC3+89	GC3+90	GC3+91	GC3+92	GC3+93	GC3+94	GC3+95	GC3+96	GC3+97	GC3+98	GC3+99	GC3+100	GC3+101	GC3+102	GC3+103	GC3+104	GC3+105	GC3+106	GC3+107	GC3+108	GC3+109	GC3+110	GC3+111	GC3+112	GC3+113	GC3+114	GC3+115	GC3+116	GC3+117	GC3+118	GC3+119	GC3+120	GC3+121	GC3+122	GC3+123	GC3+124	GC3+125	GC3+126	GC3+127	GC3+128	GC3+129	GC3+130	GC3+131	GC3+132	GC3+133	GC3+134	GC3+135	GC3+136	GC3+137	GC3+138	GC3+139	GC3+140	GC3+141	GC3+142	GC3+143	GC3+144	GC3+145	GC3+146	GC3+147	GC3+148	GC3+149	GC3+150	GC3+151	GC3+152	GC3+153	GC3+154	GC3+155	GC3+156	GC3+157	GC3+158	GC3+159	GC3+160	GC3+161	GC3+162	GC3+163	GC3+164	GC3+165	GC3+166	GC3+167	GC3+168	GC3+169	GC3+170	GC3+171	GC3+172	GC3+173	GC3+174	GC3+175	GC3+176	GC3+177	GC3+178	GC3+179	GC3+180	GC3+181	GC3+182	GC3+183	GC3+184	GC3+185	GC3+186	GC3+187	GC3+188	GC3+189	GC3+190	GC3+191	GC3+192	GC3+193	GC3+194	GC3+195	GC3+196	GC3+197	GC3+198	GC3+199	GC3+200	GC3+201	GC3+202	GC3+203	GC3+204	GC3+205	GC3+206	GC3+207	GC3+208	GC3+209	GC3+210	GC3+211	GC3+212	GC3+213	GC3+214	GC3+215	GC3+216	GC3+217	GC3+218	GC3+219	GC3+220	GC3+221	GC3+222	GC3+223	GC3+224	GC3+225	GC3+226	GC3+227	GC3+228	GC3+229	GC3+230	GC3+231	GC3+232	GC3+233	GC3+234	GC3+235	GC3+236	GC3+237	GC3+238	GC3+239	GC3+240	GC3+241	GC3+242	GC3+243	GC3+244	GC3+245	GC3+246	GC3+247	GC3+248	GC3+249	GC3+250	GC3+251	GC3+252	GC3+253	GC3+254	GC3+255	GC3+256	GC3+257	GC3+258	GC3+259	GC3+260	GC3+261	GC3+262	GC3+263	GC3+264	GC3+265	GC3+266	GC3+267	GC3+268	GC3+269	GC3+270	GC3+271	GC3+272	GC3+273	GC3+274	GC3+275	GC3+276	GC3+277	GC3+278	GC3+279	GC3+280	GC3+281	GC3+282	GC3+283	GC3+284	GC3+285	GC3+286	GC3+287	GC3+288	GC3+289	GC3+290	GC3+291	GC3+292	GC3+293	GC3+294	GC3+295	GC3+296	GC3+297	GC3+298	GC3+299	GC3+300	GC3+301	GC3+302	GC3+303	GC3+304	GC3+305	GC3+306	GC3+307	GC3+308	GC3+309	GC3+310	GC3+311	GC3+312	GC3+313	GC3+314	GC3+315	GC3+316	GC3+317	GC3+318	GC3+319	GC3+320	GC3+321	GC3+322	GC3+323	GC3+324	GC3+325	GC3+326	GC3+327	GC3+328	GC3+329	GC3+330	GC3+331	GC3+332	GC3+333	GC3+334	GC3+335	GC3+336	GC3+337	GC3+338	GC3+339	GC3+340	GC3+341	GC3+342	GC3+343	GC3+344	GC3+345	GC3+346	GC3+347	GC3+348	GC3+349</
-----------	---------	--------	----	-----	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	-----------

ALIGNMENTS

```

RESULT 1
Q723Y5 HUMAN
ID Q723Y5 HUMAN PRELIMINARY; PRT; 86 AA.
AC Q723Y5;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Rearranged VKA17 V gene segment (Fragment).
GN Name=VKA17;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RS TISSUE=Hodgkin lymphoma;
RA Tinguely M., Rosenquist R., Sundstroem C., Amini R.M., Kuppers R.,
RA Hansmann M.L., Brauning A.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC

```

Query Match	18.4%;	Score 48;	DB 2;	Length 86;
Best Local Similarity	100.0%;	Pred. No. 3.6e-37;		
Matches 48:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

RESULT 2		
Q9UL80 HUMAN		
ID	Q9UL80 HUMAN	PRELIMINARY; PRT; 114 AA.
AC	Q9UL80;	
DT	01-MAY-2000.	integrated into UniProtKB/TrEMBL.

DT 01-MAY-2000, sequence version 1.
 DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 21.
 DE Myosin-reactive immunoglobulin light chain variable region (Fragment).

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92352481; PubMed=1322670;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92352481; PubMed=1322670;
RA Stuber F., Lee S.K., Bridges S.L. Jr., Koopman W.J., Schroeder H.W. Jr.,
RA Gaskin F., Fu S.M.;
RT "A rheumatoid factor from a normal individual encoded by VH2 and V
RT kappa II gene segments."
RL Arthritis Rheum. 35:900-904(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93170387; PubMed=8436174;
RA Wagner S.D., Luzzatto L.;
RT "V kappa gene segments rearranged in chronic lymphocytic leukemia are
RT distributed over a large portion of the V kappa locus and do not show
RT somatic mutation."
RL Eur. J. Immunol. 23:391-397(1993).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92289816; PubMed=1601042;
RA Huber C., Klobeck H.G., Zachau H.G.;
RT "Ongoing V kappa-J kappa recombination after formation of a productive
RT V kappa-J kappa coding joint."
RL Eur. J. Immunol. 22:1561-1565(1992).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC
CC EMBL; AF035034; AAD56270.1; -; mRNA.
DR PIR; B49002; B49002.
DR PIR; S23638; S23638.
DR PIR; S34094; S34094.
DR PIR; S34095; S34095.
DR HSP; P01625; 11VE.
DR SMR; Q9UL80; 1-114.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS0835; IG LIKE; 1.
DR NON TER 1 114
DR NON TER 114 114
SQ SEQUENCE 114 AA; 12775 MW; 070E31E210D1CB01 CRC64;

Query Match 18.4%; Score 48; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.7e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRLIYKVNRSQVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 40 WFOQPGQSPRLIYKVNRSQVDPDRFSGSGGTDTFTLKISRVEAED 87

RESULT 3
KV2F_HUMAN
ID KV2F_HUMAN STANDARD; PRT; 133 AA.
AC P06310;
DT 01-JAN-1988, integrated into UniProtKB/Swiss-Prot.
```

```
DT 01-JAN-1988, sequence version 1.
DT 07-MAR-2006, entry version 41.
DE IG kappa chain V-II region RPMI 6410 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=86041852; PubMed=2997711;
RA Klobeck H.G., Meindl A., Combiato G., Solomon A., Zachau H.G.;
RT "Human immunoglobulin kappa light chain genes of subgroups II and
RT III."
RL Nucleic Acids Res. 13:6499-6513(1985).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; Z00020; CAA77315.1; -; Genomic_DNA.
DR PIR; A01890; K2HURP.
DR HSP; Q99M37; 1I9I.
DR SMR; P06310; 2I-133.
DR EMBL; ENSG00000173758; Homo sapiens.
DR LinkHub; P06310; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR Immunoglobulin domain; Immunoglobulin V region; Signal.
KW SIGNAL
FT CHAIN 21 133
FT REGION 21 43
FT REGION 44 59
FT REGION 60 74
FT REGION 75 81
FT REGION 82 113
FT REGION 114 122
FT REGION 123 132
FT DISULFID 43 113
FT NON TER 133
SQ SEQUENCE 133 AA; 14707 MW; 513CCAF3673009EE CRC64;

Query Match 18.4%; Score 48; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 5.4e-37;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 173 WFOQPGQSPRLIYKVNRSQVDPDRFSGSGGTDTFTLKISRVEAED 220
Db 60 WFOQPGQSPRLIYKVNRSQVDPDRFSGSGGTDTFTLKISRVEAED 107

RESULT 4
Q8TCD0_HUMAN
ID Q8TCD0_HUMAN PRELIMINARY; PRT; 239 AA.
AC Q8TCD0;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
```



```
RESULT 6
Q96QSO HUMAN
ID Q96QSO HUMAN PRELIMINARY; PRT; 159 AA.
AC Q96QSO
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DE 21-FEB-2006, entry version 20.
DE Putative matrix cell adhesion molecule-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Tilson M.D.
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY039025; AAK82649.1; -; mRNA.
DR HSP; P01869; IAE6.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG LIKE; 1.
KW Immunoglobulin domain.
SQ SEQUENCE 159 AA; 17497 MW; 5D29537E881FAF02 CRC64;

Query Match 12.3%; Score 32; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 1e-21;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTF 34
DB 20 QVQLVQSGAEVKKPGASVKVSCKASGYTF 51

RESULT 7
HV1B HUMAN
ID HV1B HUMAN STANDARD; PRT; 117 AA.
AC P01743;
DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DT 21-JUL-1986, sequence version 1.
DT 07-MAR-2006, entry version 41.
DE Ig heavy chain V-I region HG3 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=83144028; PubMed=6298778;
RA Rechavi G., Ram D., Glazer L., Zakut R., Givol D.;
RT "Evolutionary aspects of immunoglobulin heavy chain variable region
(VH) gene subgroups."
RL Proc. Natl. Acad. Sci. U.S.A. 80:855-859(1983).
CC -!- SIMILARITY: Contains 1 Ig-like (immunoglobulin-like) domain.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; J00240; AAA52988.1; -; Genomic_DNA.
DR PIR; A02024; HVH0HG.
DR HSP; P01751; INQB.
DR SMR; P01743; 20-117.
```

```
LinkHub; P01743; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG LIKE; 1.
KW Immunoglobulin domain; Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117
FT DOMAIN 20 >117
FT NON_TER 117 117
FT SEQUENCE 117 AA; 12946 MW; 2D3F92FC60CD1PE7 CRC64;

Query Match 11.1%; Score 29; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 5.3e-19;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTF 31
DB 20 QVQLVQSGAEVKKPGASVKVSCKASGYTF 48

RESULT 8
HV1G HUMAN
ID HV1G HUMAN STANDARD; PRT; 117 AA.
AC P23083;
DT 01-NOV-1991, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1991, sequence version 1.
DT 07-MAR-2006, entry version 36.
DE Ig heavy chain V-I region V35 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=88296408; PubMed=2841108;
RA Matsuda F., Lee K.H., Nakai S., Sato T., Kodaira M., Zong S.Q.,
RA Ohno H., Fukuhara S., Honjo T.;
RT "Dispersed localization of D segments in the human immunoglobulin
heavy-chain locus."
RL EMBO J. 7:1047-1051(1988).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 20-116.
RX MEDLINE=93209281; PubMed=7681398;
RA Mariette X., Tsapis A., Brouet J.C.;
RT "Nucleotide sequence analysis of the variable domains of four human
monoclonal IgM with an antibody activity to myelin-associated
glycoprotein."
RT Eur. J. Immunol. 23:846-851(1993).
CC -!- SIMILARITY: Contains 1 Ig-like (immunoglobulin-like) domain.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; X07448; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR PIR; S00476; HVH035.
DR HSP; P01751; INQB.
DR SMR; P23083; 20-117.
DR Ensembl; ENSG00000130076; Homo sapiens.
DR LinkHub; P23083; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR InterPro; IPR013106; V-set.
```


DR Pfam; PF07686; V-set; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain; Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117
FT DOMAIN 20 >117
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13009 MW; BB61CB63F8CE97BD CRC64;
Query Match 11.1%; Score 29; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 5.3e-19;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTF 31
DB 20 QVQLVQSGAEVKKPGASVKVSCKASGYTF 48

RESULT 9
Q53VP8_MOUSE
ID Q53VP8_MOUSE PRELIMINARY; PRT; 112 AA.
AC Q53VP8;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Kappa chain (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE
RX MEDLINE=86136012; PubMed=3937730;
RA Ollier P., Rocca-Serra J., Somme G., Theze J., Fougereau M.;
RT "The idiotypic network and the internal image: possible regulation of
a germ-line network by paucigene encoded Ab2 (anti-idiotypic)
antibodies in the GAT system";
RL EMBO J. 4:3681-3688 (1985).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 108-109.
RA Fougereau M.;
RL Submitted (NOV-1986) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; X03386; CAA27113.1; -; mRNA.
DR SMR; Q53VP8; 1-112.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR SMART; SM00409; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain.
FT NON_TER 1 1
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12270 MW; C844B7881A89C18A CRC64;
Query Match 10.3%; Score 27; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 194 SGVPDRFSGSGGTDFTLKISRVEAED 220
DB 61 SGVPDRFSGSGGTDFTLKISRVEAED 87

RESULT 10
KV2D_HUMAN

ID KV2D_HUMAN STANDARD; PRT; 113 AA.
AC P01617;
DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DT 21-JUL-1986, sequence version 1.
DT 07-MAR-2006, entry version 45.
DE Ig kappa chain V-II region TEW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP PROTEIN SEQUENCE (BENCE-JONES PROTEIN TEW).
RX MEDLINE=74148480; PubMed=4596149;
RA Putnam F.W., Whitley E.J. Jr., Paul C., Davidson J.N.;
RT "Amino acid sequence of a kappa Bence Jones protein from a case of
primary amyloidosis";
RL Biochemistry 12:3763-3780 (1973).
RN [2]
RP PROTEIN SEQUENCE OF 1-27 (AMYLOID PROTEIN TEW).
RX MEDLINE=73166638; PubMed=4700495;
RA Terry W.D., Page D.L., Kimura S., Isobe T., Osseman E.F.,
RA Glenner G.G.;
RT "Structural identity of Bence Jones and amyloid fibril proteins in a
patient with plasma cell dyscrasia and amyloidosis";
RL J. Clin. Invest. 52:1276-1281 (1973).
CC -!- MISCELLANEOUS: The major amyloid protein appears to be identical
CC with the Bence Jones protein isolated from the same patient.
CC -!- MISCELLANEOUS: This protein was isolated from the urine of a
CC patient with plasma cell dyscrasia and amyloidosis.
CC -!- MISCELLANEOUS: The C region of this chain has the INV (1,2)
CC marker.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR PIR; A50370; K2HUTW.
DR HSSP; Q99M37; 1191.
DR SMR; P01617; 1-111.
DR LinkHub; P01617; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR InterPro; IPR013106; V-Set.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00409; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Amyloid; Bence-Jones protein; Direct protein sequencing;
KW Immunoglobulin domain; Immunoglobulin V region.
FT CHAIN 1 >113
FT CHAIN 1 113
FT REGION 1 23
FT REGION 24 39
FT REGION 40 54
FT REGION 55 61
FT REGION 62 93
FT REGION 94 102
FT REGION 103 112
FT REGION 103 112
FT DISULFID 23 93
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12316 MW; 0C3C38F81F1843CA CRC64;
Query Match 10.3%; Score 27; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY .194 SGVPDRFSGSGGTDFTLKISRVEAED 220
DB 61 SGVPDRFSGSGGTDFTLKISRVEAED 87

```

RESULT 11
KV2G MOUSE STANDARD; PRT; 113 AA.
AC P01631;
DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DT 21-JUL-1986, sequence version 1.
DE 07-MAR-2006, entry version 39.
DE Ig kappa chain V-II region 26-10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=83178921; PubMed=6404298;
RA Novotny J., Margolies M.N.;
RT "Amino acid sequence of the light chain variable region from a mouse
RL anti-digoxin hybridoma antibody."
RL Biochemistry 22:1153-1158(1983).
CC -!- MISCELLANEOUS: This chain was isolated from an IgG2a hybridoma
CC protein that binds digoxin.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR PIR: A01914; KVM526.
DR HSP; Q99M37; I191.
DR Ensembl; ENSMUSG00000055315; Mus musculus.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG V.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Immunoglobulin V region; Monoclonal antibody.
KW Immunoglobulin V region; Hybridoma; Immunoglobulin domain;
FT CHAIN 1 >113
FT REGION 1 23
FT REGION 24 39
FT REGION 40 54
FT REGION 55 61
FT REGION 62 93
FT REGION 94 102
FT REGION 103 112
FT REGION 123 93
FT DISULFID 23 93
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12273 MW; F9P39CE94A84C2A CRC64;

Query Match 10.3%; Score 27; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 SGVPRFSGSGGTDTLTKISRVEAD 220
Db 61 SGVPRFSGSGGTDTLTKISRVEAD 87

RESULT 12
Q5F210 MOUSE PRELIMINARY; PRT; 115 AA.
AC Q5F210;
DT 15-MAR-2005, integrated into UniProtKB/TrEMBL.
DT 15-MAR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Kappa light chain variable region (Fragment).
DE Name=IgG1 anti-TS1 VL;

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22716456; PubMed=12833571; DOI=10.1002/jmr.617;
RA Erlandsson A., Holm P., Ullen A., Stigbrand T., Sundstrom B.E.;
RT "Studies of the interactions between the anticytokerin 8 monoclonal
RT antibody TS1, its antigen and its anti-idiotypic antibody alphaTS1."
RL J. Mol. Recognit. 16:157-163(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Erlandsson A.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ884575; CA156337.1; -; mRNA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG V.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Immunoglobulin domain.
FT NON_TER 1 115
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12560 MW; E4D3BF3D63E88007 CRC64;

Query Match 10.3%; Score 27; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 4.2e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 SGVPRFSGSGGTDTLTKISRVEAD 220
Db 61 SGVPRFSGSGGTDTLTKISRVEAD 87

RESULT 13
KV2E HUMAN STANDARD; PRT; 117 AA.
AC P06309;
DT 01-JAN-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-JAN-1988, sequence version 1.
DT 07-MAR-2006, entry version 43.
DE Ig kappa chain V-II region GM607 precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=84191506; PubMed=6325927;
RA Klobeck H.G., Solomon A., Zachau H.G.;
RT "Contribution of human V kappa II germ-line genes to light-chain
RT diversity."
RL Nature 309:73-76(1984).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; Z00009; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR PIR; A01889; K2HUGM.
DR HSP; Q99M37; I191.
DR SMR; P06309; 5-115.
DR LinkHub; P06309; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.

```

```
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Immunoglobulin V region; Signal.
FT SIGNAL <1 4
FT CHAIN 5 117 Ig kappa chain V-II region GM607.
FT /FTID=PRO 0000015172.
FT REGION 5 27 Framework-1.
FT REGION 28 43 Complementarity-determining-1.
FT REGION 44 58 Framework-2.
FT REGION 59 65 Complementarity-determining-2.
FT REGION 66 97 Framework-3.
FT REGION 98 106 Complementarity-determining-3.
FT REGION 107 116 Framework-4.
FT DISULFID 27 97 By similarity.
FT NON_TER 1 1
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12664 MW; 92C57DC719B558B1 CRC64;

Query Match 10.3%; Score 27; DB 1; Length 117;
Best Local Similarity 100.0%; Pred.No. 4.2e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 SGVPRFSGSGGTDFTLKISRVEAD 220
Db 65 SGVPRFSGSGGTDFTLKISRVEAD 91

RESULT 14
Q65ZC0 MOUSE PRELIMINARY; PRT; 219 AA.
AC Q65ZC0;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Kappa light chain C_region (Fragment).
GN Name=Igk-C;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Balb/c; TISSUE=Spleen.
RX MEDLINE=96319505; PubMed=8768802;
RA Kipp B., Schlaak M., Becker W.M.;
RT "Cloning and expression of a recombinant mouse Fab-fragment
RT recognizing a defined linear epitope of Chironomus thummi major
RT allergen Chi t 1."
RL Int. Arch. Allergy Immunol. 110:348-353 (1996).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; Z37499; CAA85724.1; -; mRNA.
DR MGI; MGI:96495; Igk-C.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig.cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07654; Cl-set; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00407; IGL1; 1.
DR SMART; SM00406; IGV; 1.

DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Repeat.
FT NON_TER 1 1
FT NON_TER 219 219
SQ SEQUENCE 219 AA; 23945 MW; 7E1B82A14EAF8445 CRC64;

Query Match 10.3%; Score 27; DB 2; Length 219;
Best Local Similarity 100.0%; Pred.No. 7.6e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 SGVPRFSGSGGTDFTLKISRVEAD 220
Db 61 SGVPRFSGSGGTDFTLKISRVEAD 87

RESULT 15
Q58E08 MOUSE PRELIMINARY; PRT; 239 AA.
AC Q58E08;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Igk-C protein.
GN Name=Igk-C;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung.
RC MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer.;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung.
RC MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer.;
RG NIH MGC Project;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; BC091750; AAH91750.1; -; mRNA.
DR SMR; Q58E08; 21-239.
DR MGI; MGI:96495; Igk-C.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig.cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig_v.
```


XX New scFv Bc-12 or Bc-15 antibody comprises ligand mimetic, which
PT specifically binds to activated alphavbeta3 integrin receptor, useful for
PT treating, e.g. neoplastic disease, solid tumor, leukemia, or benign or
PT malignant breast cancer.
XX
PS Claim 2; SEQ ID NO 2; 109pp; English.
XX
CC The invention describes an antibody comprising a ligand mimetic, which
CC specifically binds to an activated alphavbeta 3 integrin receptor
CC differentially produced on a cell in a metastatic state compared to a
CC similar, non-metastatic cell, and does not bind to a non-activated
CC alphavbeta 3 integrin receptor. Also described are: a pharmaceutical
CC composition comprising the antibody; treating a disease state in a mammal
CC ; a cell line comprising a tumor cell variant with a metastatic homing
CC propensity to a target tissue; producing an antibody phage population
CC having affinity for a tumor cell target; detecting tumor cells in a
CC mammal by treatment with a cancer therapeutic; inducing or enhancing an
CC immune response to an antigen in a mammal; detecting an activated cell
CC surface receptor on a metastatic tumor cell surface in a mammalian tissue
CC sample; interfering with cells liable to undergo metastasis associated
CC with a disease state; identifying cells liable to undergo metastasis
CC associated with a disease state; an isolated Bc-12 or Bc-15
CC polynucleotide comprising a nucleotide sequence that has at least 90%
CC percent identity to a fully defined 802 or 752 bp sequence (SEQ ID NO. 1
CC or 3) given in the specification; an isolated polypeptide comprising a
CC nucleotide sequence that has at least 90% sequence identity to SEQ ID NO.
CC 1 or 3, or shares a biological function with Bc-12 or Bc-15; a vector
CC comprising the polynucleotide; an expression vector comprising the
CC polynucleotide in which the nucleotide sequence of the polynucleotide is
CC operatively linked with a regulatory sequence that controls expression of
CC the polynucleotide in a host cell; a host cell comprising the
CC polynucleotide or progeny of the cell; and determining anti-metastatic
CC activity of a test compound in a mammal. The antibody is useful for
CC treating neoplastic disease, solid tumor, hematological malignancy,
CC leukemia, colorectal cancer, benign or malignant breast cancer, uterine
CC cancer, uterine leiomyomas, ovarian cancer, endometrial cancer,
CC polycystic ovary syndrome, endometrial polyps, prostate cancer, prostatic
CC hypertrophy, pituitary cancer, adenomyosis, adenocarcinomas, meningioma,
CC melanoma, bone cancer, multiple myeloma, CNS cancer, glioma, or
CC astroblastoma. This is the amino acid sequence of single chain variable
CC fragment antibody Bc-12.
XX
SQ Sequence 260 AA;

Query Match 98.8%; Score 1381; DB 9; Length 260;
Best Local Similarity 99.6%; Pred. No. 3.2e-91;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGITWVRQAPGQGLEWMGWINNGNTHY 60
Db 2 MAQVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGITWVRQAPGQGLEWMGWINNGNTHY 61

Qy 61 AQKPFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDDPYVWGQGLTVTVSSGG 120
Db 62 YQKPFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDDPYVWGQGLTVTVSSGG 121

Qy 121 GSGGGGGGGGSEIVLTQSLPLVTLGQPASISCRSSQNLVYSDGNTYLSWFOORPQG 180
Db 122 GSGGGGGGGGSEIVLTQSLPLVTLGQPASISCRSSQNLVYSDGNTYLSWFOORPQG 181

Qy 181 SPRRLIYKVNDRSGVPRDFSGSGSGTFTLKI SRVEAEDIGVYCMQGTWPPPTFGQ 240
Db 182 SPRRLIYKVNDRSGVPRDFSGSGSGTFTLKI SRVEAEDIGVYCMQGTWPPPTFGQ 241

Qy 241 TKVEIKRGLGLVDYKDD 259
Db 242 TKVEIKRGLGLVDYKDD 260

RESULT 2
ID ABP45953

ABP45953 standard; protein; 253 AA.

XX ABP45953;
AC 19-AUG-2002 (first entry)
XX
DE Human Blys binding scFv SEQ ID 1964.
XX
KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200202641-A1.
XX
PD 10-JAN-2002.
XX
XX 15-JUN-2001; 2001WO-US019110.
XX
PR 16-JUN-2000; 2000US-0212210P.
PR 17-OCT-2000; 2000US-0240816P.
PR 16-MAR-2001; 2001US-0276248P.
PR 21-MAR-2001; 2001US-0277379P.
PR 25-MAY-2001; 2001US-0293499P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX WPI; 2002-114799/15.
XX
XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
XX diagnosis and treatment of cancers and immune disorders.
XX
XX Claim 1; Page 2744-2745; 3148pp; English.
XX
CC This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic, and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of Blys. The antibodies bind to Blys
CC and so may be used to detect and quantitate the presence of Blys in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of Blys. They may also be
CC administered to treat diseases associated with aberrant Blys expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID), and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method of
CC the invention
XX
SQ Sequence 253 AA;

Query Match 81.2%; Score 1135; DB 5; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.4e-73;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNTHY 60

Qy 61 AQKPFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDDPYVWGQGLTV 115
Db 61 AQKPFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDDPYVWGQGLTV 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSLPLVTLGQPASISCRSSQNLVYSDGNTYLSW 173

Db	121	VSSGGSGSGSGGSGGSGALDVMTTQSLSPVTLLQGQPASISCRSQSLVSDGNTYLNW	189
Qy	174	FQQRFGQSPRLIYKVSNRDSCVPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQTTHWP	233
Db	181	FQQRFGQSPRLIYKVSNRDSCVPDRFSGSGSGTDFTLKISRVEAEDVGYYCMQATRW-	239
Qy	234	PRTFGQGTKEIKR	247
Db	240	PFTFGQGTKEIKR	253
 RESULT 3 ADG96780 standard; protein; 253 AA. XX AC ADG96780; XX AC XX DT 11-MAR-2004 (first entry) XX XX Single chain antibody that immunospecifically binds Blys SeqID 1964. XX KW antibody; B lymphocyte stimulator; Blys; tumour necrosis factor; KW B cell proliferation; differentiation; scFv; myasthenia gravis; KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia; KW carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective; KW antiinflammatory; antiasthmatic; antiallergic; cytostatic. XX KS Unidentified. XX OS XX PN WO2003055979-A2. XX PD 10-JUL-2003. XX PF 14-NOV-2002; 2002WO-US036496. XX PR 16-NOV-2001; 2001US-0331469P. XX PR 19-DEC-2001; 2001US-0340817P. XX PA (HUMA-) HUMAN GENOME SCI INC. XX PI Ruben SM, Barash SC, Choi GH, Vaughan TU, Hilbert D; XX WPI; 2003-505530/47. XX XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator PT (Blys), useful for detecting and treating diseases or disorders e.g. PT rheumatoid arthritis, asthma and leukemia. XX XX Example 1; SEQ ID NO 1964; 394pp; English. XX CC This invention relates to novel antibodies that immunospecifically bind CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to CC chromosome 13q34 and encodes a protein that is a member of the tumour CC necrosis factor superfamily and induces both in vivo and in vitro B cell CC proliferation and differentiation. Specifically, it refers to single CC chain antibody molecules (scFvs) derived, preferably, from the variable CC heavy CDR3 region that immunospecifically bind to a polypeptide, or CC fragment thereof, of either human, murine, rat or monkey Blys. The CC present invention refers to the use of such antibodies in various methods CC for the detection, diagnosis and prognosis of diseases related to the CC aberrant expression or inappropriate function of Blys or its receptor. As CC such, these compositions are useful for identifying immune disorders CC including myasthenia gravis and multiple sclerosis, inflammatory CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such CC as AIDS and proliferative disorders including leukaemia, carcinoma and CC lymphoma. Accordingly, they can be described as exhibiting various CC activities such as antirheumatic, antiarthritic, neuroprotective, CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This CC polypeptide sequence is a single chain antibody that binds Blys of the CC invention. NOTE: The sequence data for this patent did not form part of CC the printed specification, but was obtained in electronic format CC directly from WIPO at ftp.wipo.int/pub/published pct sequences.			

XX	Sequence 253 AA;
SQ	
Query Match	81.2%; Score 1135; DB 7; Length 253;
Best Local Similarity	86.2%; Pred. No. 1.4e-73;
Matches 219; Conservative	15; Mismatches 10; Indels 10; Gaps 4
Qy	3 QVQLVSGAEEVKPGASVKYSCKASGYTFSNYGITWVRQAPOGLLEWMGWIN--NGNTHY 60
Dd	1 QVQLQQSGAEVKPEPGASVKVSCKASGYTFTSYGISWVRQAPOGLLEWMGWISAYNGNTNY 60
Qy	61 AOKPQGRTVTMTDTSTAYMELSRLSDDDTAIVYCAEDPRGDDDE----PYWGOGTLVT 115
Dd	61 AQKLQGRVTMTDTSTAYMELSRLSDDDTAIVYCARNPYYDSSEGFDFYMGOGTWMT 120
Qy	116 VSSGGGSGGGSGGGSGGS--EIVLTQPSLPLSVTLGPASISCRSSQNLVYSDGNLYLSW 173
Dd	121 VSSGGGSGGGSGGGSGGSALDVMTQPSLPLSVTLGPASISCRSSQSLVYSDGNLYLNW 180
Qy	174 FQRPQGSPRRLIYKVNRSRGVPDRFSGSGGTDFTLKISRVEADIGVVYCMQGTHTWP 233
Dd	181 FQRPQGSPRRLIYKVNRSRGVPDRFSGSGGTDFTLKISRVEADVGVVYCMQATRW- 239
Qy	234 PRTFGQGTKVEIKR 247
Dd	240 PPTFGQGTKMEIKR 253
RESULT 4	
AED78833	ID AED78833 standard; protein; 253 AA.
XX	AED78833;
AC	
XX	12-JAN-2006 (first entry)
DT	
XX	Human B Lymphocyte binding scFv fragment protein, SEQ ID 1964.
DE	
DE	
XX	
KW	Antiinflammatory; Dermatological; Immunosuppressive; Antirheumatic;
KW	Antiarthritic; Neuroprotective; Muscular-Gen.; Antiasthmatic;
KW	Antiallergic; Antimicrobial; Anti-HIV; Cytostatic; B-lymphocyte;
KW	antibody; autoimmune disease; b-cell lymphoma;
KW	systemic lupus erythematosus; rheumatoid arthritis; immune disorder;
KW	inflammation; infectious disease; hyperproliferation.
OS	Homo sapiens.
XX	
XX	US2005255532-A1.
PN	
XX	17-NOV-2005.
XX	
PF	10-FEB-2005; 2005US-00054515.
XX	
XX	16-JUN-2000; 2000US-0212210P.
PR	17-OCT-2000; 2000US-0240816P.
PR	16-MAR-2001; 2001US-0276248P.
PR	21-MAR-2001; 2001US-0277379P.
PR	25-MAY-2001; 2001US-0293499P.
PR	18-JUN-2001; 2001US-00880748.
PR	16-DEC-2001; 2001US-0331469P.
PR	19-NOV-2001; 2001US-0340817P.
PR	14-NOV-2002; 2002US-00293418.
PR	11-FEB-2004; 2004US-0543296P.
PR	18-JUN-2004; 2004US-0580347P.
XX	
PA	(RUBE/) RUBEN S M.
PA	(BARA/) BARASH S C.
PA	(CHOI/) CHOI G H.
PA	(VAUG/) VAUGHAN T.
PA	(HILB/) HILBERT D.
XX	
XX	Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
PI	
XX	

DR WPI; 2005-808635/82.
 XX New antibodies that immunospecifically binds to B Lymphocyte Stimulator
 PT protein, useful for diagnosing, treating, or preventing autoimmune
 PT disease, e.g. systemic lupus erythematosus or rheumatoid arthritis, or B
 PT cell cancer.
 XX
 PS Claim 1; SEQ ID NO 1964; 240pp; English.
 XX
 CC The invention relates to a novel antibody that immunospecifically binds
 CC to a B Lymphocyte Stimulator protein. The protein comprises an amino acid
 CC sequence that is 85% identical to the VH domain of any one of the single-
 CC chain variable fragments (scFvs) of SEQ ID NOS. 1-2128, and/or an amino
 CC acid sequence that is at least 85% identical to the VL domain of any one
 CC of the scFvs of SEQ ID NOS. 1-2128. The invention further comprises: an
 CC isolated nucleic acid molecule encoding the antibody; an isolated cell
 CC line that expresses the antibody; a method for detecting the expression
 CC of a B Lymphocyte Stimulator protein; a method for diagnosing an
 CC autoimmune disease or a B cell cancer; and a method for treating,
 CC preventing, or ameliorating an autoimmune disease or a B cell cancer. The
 CC antibody is useful for detecting expression of B Lymphocyte Stimulator
 CC protein, and in diagnosing, treating, preventing, or ameliorating an
 CC autoimmune disease or a B cell cancer. The autoimmune disease is systemic
 CC lupus erythematosus or rheumatoid arthritis. It can also be used for
 CC diagnosing, treating, and preventing immune disorders (e.g. multiple
 CC sclerosis, myasthenia gravis, or Hashimoto's disease), inflammatory
 CC disorders (e.g. asthma or allergic disorders), infectious diseases (e.g.
 CC AIDS), and proliferative disorders (e.g. leukemia, carcinoma, or
 CC lymphoma). This sequence represents a single-chain variable fragment
 CC polypeptide that immunospecifically binds to a B Lymphocyte Stimulator
 CC protein of the invention. Note: This sequence is not shown in the
 CC specification. It has been electronically downloaded from the USPTO
 CC website.
 XX
 SQ Sequence 253 AA;

Query Match 81.2%; Score 1135; DB 9; Length 253;
 Best Local Similarity 86.2%; Pred. No. 1.4e-73;
 Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
 QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGTSVVRQAPQGQLEWGWGWIN--NGNTHY 60
 DB 1 QVQLQSGAEVKKPGASVKVSCKASGYTFTSYGTSVVRQAPQGQLEWGWGWIN--NGNTHY 60
 QY 61 AOKQGRVTMTTDTSTAYMELSLRSDDTAVYVCARDPRGDDDE----PWGQGTAVT 115
 DB 61 AOKLQGRVTMTTDTSTAYMELSLRSDDTAVYVCARNPYDSSGFFDYWGQGTAVT 120
 QY 116 VSSGGSGGGSGGGSGGG--EIVLTQSPSLPVLTPGQASISCRSSQNLVYSDGNTYLSW 173
 DB 121 VSSGGSGGGSGGGSGGG--EIVLTQSPSLPVLTPGQASISCRSSQNLVYSDGNTYLSW 180
 QY 174 FQQRPGQSPRLIYKVSNRDGVDPFRFGSGSGTDFTLKISRVEAEDIGVYVCMQGTWHP 233
 DB 181 FQQRPGQSPRLIYKVSNRDGVDPFRFGSGSGTDFTLKISRVEAEDIGVYVCMQGTWHP 239
 QY 234 PRTFGQGTWHP 247
 DB 240 PFTFGQGTWHP 253

RESULT 5
 AAY21884
 ID AAY21884 standard; protein; 242 AA.
 XX
 AC AAY21884;
 XX
 DT 24-SEP-1999 (first entry)
 XX
 DE Amino acid sequence of antibody BIOC7.
 XX
 KW prostate-specific antigen; PSA; anti-chymotrypsin; ACT; antibody;
 KW diagnostic assay; PSA-ACT; prostate cancer; tumour; immunotherapy; BIOC7.

XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 102 /label= unknown
 FT /note= "encoded by AST"
 FT Misc-difference 114 /label= unknown
 FT /note= "encoded by TCN"
 FT Misc-difference 155 /label= unknown
 FT /note= "encoded by TGT"
 PN EP934953-A2.
 PD 11-AUG-1999.
 XX 01-DEC-1998; 98EP-00122546.
 XX 03-DEC-1997; 97US-0067428P.
 XX (BOFF) BOEHRINGER MANNHEIM CORP.
 XX Winter GP, Mahoney W, Sawyer JR;
 DR WPI; 1999-432068/37.
 DR N-PSDB; AAX86942.
 XX New anti-complex antibody useful for diagnosing prostate cancer.
 XX Claim 29; Page 28-30; 42pp; English.
 CC The invention relates to an antibody that binds a complex between
 CC prostate-specific antigen (PSA) and anti-chymotrypsin (ACT), and has an
 CC affinity for the complex which is at least 10 fold higher than the
 CC affinity for either PSA or ACT alone. The antibody is used in diagnostic
 CC assays to detect PSA-ACT in serum samples from patients. The levels of
 CC PSA-ACT complex increase in patients suffering from prostate cancer
 CC compared to the levels in patients with benign or no growths in the
 CC prostate. Therefore detection of PSA in complex with ACT is useful for
 CC the early detection of prostate tumours, by distinguishing between benign
 CC and malignant conditions of the prostate as well as for the management of
 CC patients with prostate cancer, such as the disclosure of metastasis and
 CC the monitoring of the PSA levels after treatment. The antibodies may also
 CC be used in immunotherapy, affinity chromatography and isolation or
 CC purification of PSA-ACT. Unlike prior art antibodies which bind to PSA-
 CC ACT complexes but may also bind PSA or ACT alone, the present antibody is
 CC specific for PSA-ACT alone. Diagnostic assays using the antibodies are
 CC more accurate in diagnosing prostate cancer as they only detect the
 CC intact complex of PSA-ACT. Sequences AAY21880-884 represent specific
 CC examples of antibodies directed against PSA-ACT. The present sequence
 CC represents the amino acid sequence of antibody BIOC7
 XX Sequence 242 AA;

Query Match 72.4%; Score 1011.5; DB 2; Length 242;
 Best Local Similarity 78.7%; Pred. No. 1e-64;
 Matches 196; Conservative 20; Mismatches 24; Indels 9; Gaps 3;
 QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGTSVVRQAPQGQLEWGWGWIN--NGNT 58
 DB 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGTSVVRQAPQGQLEWGWGWIN--NGNT 60
 QY 59 HYAQKFGQGRVTMTTDTSTAYMELSLRSDDTAVYVCARDPRGDDDEPWGQGTAVTSS 118
 DB 61 KYSQKFGQGRVTMTTDTSTAYMELSLRSDDTAVYVCARGXR-----FWGQGTAVTSS 115
 QY 119 GGGSGGGSGGGSGGGSEIVLTQSPSLPVLTPGQASISCRSSQNLVYSDGNTYLSWFOQRP 178
 DB 116 GGGSGGGSGGGSGGGSEIVLTQSPSLPVLTPGQASISCRSSQNLVYSDGNTYLSWFOQRP 175
 QY 179 GQSPRLIYKVSNRDGVDPFRFGSGSGTDFTLKISRVEAEDIGVYVCMQGTWHPPTTFG 238

Db 176 GQPQLLIYVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAEDVGYVYCMQSIQL--HSPG 233
 Qy 239 QGTQVEIKR 247
 Db 234 QGTQVEIKR 242
 RESULT 6
 ADP03811
 ID ADP03811 standard; protein; 286 AA.
 XX AC ADP03811;
 DT 26-AUG-2004 (first entry)
 XX Human antibody related protein sequence, SEQ ID 184.
 DE Cytostatic; Antiarteriosclerotic; Antiinflammatory; Immunosuppressive;
 KW human; antibody; cancer lesion; arteriosclerosis; inflammatory disease;
 KW autoimmune disease; cancer.
 XX OS Homo sapiens.
 XX PN WO2004048571-A1.
 XX PD 10-JUN-2004.
 XX PF 21-NOV-2003; 2003WO-JP014919.
 XX PR 22-NOV-2002; 2002JP-00339241.
 XX (CHUS) CHUGAI SEIYAKU KK.
 PA (PHAR-) PHARMALOGICALS RES PTE LTD.
 XX Tsuchiya M, Suzuki M, Yoshida K, Fujii E, Mateubara K, Tsunoda H;
 DR NPI; 2004-450382/42.
 DR N-PSDB; ADP03810.
 XX Isolating polynucleotide that encodes antibody which acts against
 PT lesioned tissue, involves isolating B cells that is infiltrated into
 PT lesioned tissue, and acquiring polynucleotide that encodes antibody from
 PT isolated B cells.
 XX Example 4; SEQ ID NO 184; 200pp; Japanese.
 XX The present invention relates to novel antibody sequences, which acts
 CC against lesioned tissue. Also claimed is a method (M1) for isolating
 CC polynucleotide encoding the antibodies, which involves (a) isolating B
 CC cells that is infiltrated into lesioned tissue, and (b) acquiring
 CC polynucleotide that encodes an antibody from the isolated B cells. The
 CC antibodies are useful for treating cancer lesions, arteriosclerosis,
 CC inflammatory disease or autoimmune disease. The present sequence was used
 CC to illustrate the invention.
 XX SQ Sequence 286 AA;
 Query Match 70.4%; Score 983.5; DB 8; Length 286;
 Best Local Similarity 72.0%; Pred. No. 1.2e-62;
 Matches 198; Conservative 21; Mismatches 31; Indels 25; Gaps 6;
 Qy 1 MAQVQLVQSGAEVKKPKASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNT 58
 Db 23 MAQVQLVQSGAEVKKPKASVKVCKASGYTFTGYTHHWVRQAPGQGLEWMGWINPNSGGT 82
 Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELSLRSDDTAVYVCAR-----DPRGDDEPY 107
 Db 83 KYAQKFGQRTVMTTDTSTSTAYMELSLRSDDTAVYVCARGYDILTYGWFDP----- 135
 Qy 108 WGGGTLVTSSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 167
 Db 136 WGGGTLVTSSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 195

Qy 168 N-TYLSWFOQRQSQPRRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAEDIGVYVC 226
 Db 196 NKNYLAWYQQKFGQPPKLLIYWASTRESGVDPDRFSGSGGTDTFTLTSLQAEADVAVYVC 255
 Qy 227 MQGTHWPPRTFGQTKVEIKRGLGLGVLDYKDDDDK 261
 Db 256 QQYYSTPP-TFGQTKVEIKR---REFDYKDDDDK 286
 RESULT 7
 AAB70171
 ID AAB70171 standard; protein; 251 AA.
 XX AC AAB70171;
 DT 26-APR-2001 (first entry)
 XX Recombinant antibody #1.
 KW Antibody; sperm; S19; contraception.
 XX OS Synthetic.
 XX PN WO200107083-A1.
 XX PD 01-FEB-2001.
 XX PF 21-JUL-2000; 2000WO-US019843.
 XX PR 23-JUL-1999; 99US-0145512P.
 XX (UYVI-) UNIV VIRGINIA PATENT FOUND.
 PI Herr JC, Norton EJ, Diekman AB;
 XX WPI; 2001-182730/18.
 DR New recombinant antibody derivative of monoclonal antibody S19, useful as
 PT a reagent for purifying or detecting human spermatozoa, as an active
 PT ingredient of a spermistatic agent, or as a component of a spermicidal
 PT contraceptive.
 XX Claim 21; Page 41-42; 48pp; English.
 XX The present invention relates to a recombinant antibody capable of
 CC specifically binding to sperm agglutination antigen-1 (SAGA-1). The
 CC recombinant antibody has two peptide fragments of the S19 antibody and
 CC the fragments are joined together by a linker. The recombinant monoclonal
 CC antibodies are useful in a passive immunity composition for contraception
 CC as they inhibit the ability of sperm to fertilize an egg. The antibodies
 CC may be used e.g. as an active ingredient of a spermistatic agent, or as a
 CC component of a spermicidal contraceptive.
 XX SQ Sequence 251 AA;
 Query Match 70.3%; Score 982.5; DB 4; Length 251;
 Best Local Similarity 77.1%; Pred. No. 1.3e-62;
 Matches 192; Conservative 24; Mismatches 30; Indels 3; Gaps 2;
 Qy 1 MAQVQLVQSGAEVKKPKASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNT 58
 Db 1 MAQVQLVQSGAEVKKPKASVKVCKASGYTFTTYMHVVRQAPGQGLEWIGDIYPSGSDS 60
 Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELSLRSDDTAVYVCARDPRDDEPYWGGGTLVTSS 118
 Db 61 NYDVKFNRTITADTSTSTAYMQLSLRSEDTAVYVCARGDYGCPFYVWGGGTLVTSS 120
 Qy 119 GGGSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 178
 Db 121 GGGSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 180
 Qy 179 GQSPRLIYKVSNRDSGVDPDRFSGSGGTDTFTLKISRVEAEDIGVYVCNQGTHWPPRTFG 238

Db 181 GQSPQLIYRVSNGPVPDRFSGSGGTDFTLKISRVEADGVGYCSQSH-VPFTFG 239
 QY 239 QGTVKEIKR 247
 Db 240 QGTVKEIKR 248

RESULT 8
 AAU90900
 ID AAU90900 standard; protein; 247 AA.
 XX AAU90900;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Insulin/insulin-like growth factor receptor-binding peptide #2856.
 XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
 KW ophthalmological; insulin; receptor; gene therapy; diabetes;
 KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
 KW diabetic retinopathy; neurological diseases; stroke; diabetic neuropathy.
 XX
 OS Synthetic.
 XX
 XX WO200172771-A2.
 XX
 XX
 PD 04-OCT-2001.
 XX
 PF 29-MAR-2000; 2000WO-US008528.
 XX
 XX 29-MAR-2000; 2000WO-US008528.
 XX
 XX (DGIB-) DGI BIOTECHNOLOGIES LLC.
 XX (NOVO) NOVO NORDISK AS.
 XX
 PI Beasley J, Blume AJ, Schaeffer L, Pillula R, Brandt J;
 PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandecki WS;
 PI Hansen PH, Ravera M, Hsiao K;
 XX WPI; 2002-025774/03.
 DR
 XX Modulating insulin activity in mammalian cells, for treating e.g.
 PT diabetes and tumors, comprises using peptides that bind to insulin or
 PT insulin-like growth factor receptors.
 XX
 XX Example 5; Fig 34; 390pp; English.
 XX
 CC The invention relates to a method of modulating insulin activity in
 CC mammalian cells by administering a peptide that binds the insulin
 CC receptor (IR). A composition containing a peptide, optionally expressed
 CC from gene therapy vectors, that binds to Site 1 of IR and an insulin
 CC agonist are useful for treating diabetes. Also, peptides that are
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
 CC receptor agonists are useful for treating neurological diseases.
 CC including stroke and diabetic neuropathy. The peptides are also useful in
 CC screening for compounds that bind to IR or IGF-1 receptor, potential
 CC therapeutics and research reagents. AAU8034-AAU90957 represent IR and/or
 CC IGF-1 receptor-binding peptides and related amino acid sequences of the
 CC invention
 XX
 XX Sequence 247 AA;
 XX
 CC Query Match 70.0%; Score 978.5; DB 5; Length 247;
 CC Best Local Similarity 74.2%; Pred. No. 2.4e-62;
 CC Matches 184; Conservative 28; Mismatches 31; Indels 5; Gaps 2;
 XX
 QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNVTWVQAPCGGLEWGWL-----NNGNT 58
 Db 1 QVQLVESGGGLVKGPGSLRLSCAASGFTFSNAMSWSVRQAPGKGLIEWGRINSKTDGTT 60

QY 59 HYAQFQGRVTMTTSTSTAYMELRSRSDDTAVYVCARDPRGDDPEYWGQTLTVTSS 118
 Db 61 DYAAPVKGRFTISRDDSKNTLYLQMSLAKTEDAVYVCTVALSADRGWQGLTVTSS 120
 QY 119 GGGSGGGSGGGSGSEIVLTQSPLSI.PVTLGQPASTSCSSONLYVSDGNTYLSWFQORP 178
 Db 121 GGGSGGGSGGGSGGSDVVMVTQSPLSI.PVTLGQPASTSCSSQSLVYSDGNTYLNWFQORP 180
 QY 179 GQSPRLIYKVSNRDVGVPDRFSGSGGTDFTLKISRVEADGVGYCMQTHWPPRTFG 238
 Db 181 GQSPRLIYKVSNRDVGVPDRFSGSGGTDFTLKISRVEADGVGYCMQTHW-PPYTFG 239
 QY 239 QGTVKEIK 246
 Db 240 QGTVKEIK 247

RESULT 9
 AEA44228
 ID AEA44228 standard; protein; 524 AA.
 XX AEA44228;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 XX Anti-TPO single chain antibody hVB22B q-wz5.
 DE
 DE hemostatic; single chain antibody; protein purification;
 KW antibody engineering; antibody production; thrombocytopenia; hemostatic;
 KW hematological disease; hVB22B q-wz5.
 XX
 OS Homo sapiens.
 XX
 XX WO2005056604-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 10-DEC-2004; 2004WO-JP018506.
 XX
 PR 12-DEC-2003; 2003JP-00415746.
 PR 12-MAR-2004; 2004JP-00071763.
 PR 27-AUG-2004; 2004JP-00248323.
 XX
 XX (CHUS) CHUGAI SEIYAKU KK.
 XX
 XX Tsunoda H, Nakano K, Orita T, Tsuchiya M, Hirata Y;
 PI
 PI WPI; 2005-445149/45.
 DR N-PSDB; AEA44227.
 XX
 XX Anti-Mpl antibodies including antihuman Mpl antibody and variants like
 PT genetically-engineered antihuman Mpl diabody, applicable in drug
 PT compositions for use in chemotherapy of thrombocytopenia and other
 PT diseases associated with Mpl mutation.
 XX
 PS Claim 29; SEQ ID NO 293; 106pp; Japanese.
 XX
 CC The invention describes an antibody is a single-chain polypeptide with
 CC TPO (thrombopoietin) receptor (Mpl)-binding activity that contains the 2
 CC heavy-chain variable regions and 2 light-chain variable regions. Also
 CC described are: a chimeric antibody for binding with Mpl; an antibody for
 CC binding with soluble Mpl; antibodies for binding with human Mpl and
 CC monkey Mpl; antibodies with agonistic activity against human Mpl and
 CC monkey Mpl; antibodies with binding activity for soluble Mpl; antibodies
 CC with TPO agonistic activity; antibodies containing a heavy-chain variable
 CC region with CDR (complementary-determinant regions 1, 2 and 3);
 CC antibodies containing a light-chain variable region with CDR 1, 2 and 3;
 CC antibodies containing heavy and light-chain variable regions with CDR 1,
 CC 2 and 3 selected from the already-specified groups in 18 combinations;
 CC antibodies containing a heavy-chain variable region with FRL 1, 2, 3 and 4;
 CC antibodies containing a light-chain variable region FRL 1, 2, 3 and 4;
 CC epitope-recognizing antibodies that can recognize the antibodies with the
 CC heavy and light-chain variable regions and amino acid sequences;

CC antibodies that can recognize the an amino acid moiety in human Mpl from
 CC positions 26-274; polynucleotides encoding these antibodies;
 CC polynucleotides hybridizable with the polynucleotides that encode the
 CC antibodies under stringent conditions and the antibodies encoded by which
 CC have equivalent activity as such antibodies; vectors containing these
 CC polynucleotides; host cells carrying such polynucleotides or vectors; and
 CC drug compositions containing these antibodies. The antibodies are
 CC applicable in drug compositions for use in chemotherapy of
 CC thrombocytopenia and other diseases associated with Mpl mutation such as
 CC congenital amegakaryocytic thrombocytopenia. These antibodies can include
 CC anti-human Mpl antibody and variants like genetically-engineered anti-human
 CC Mpl diabody and (humanized) anti-human Mpl sc(Fv)2. Such antibodies are
 CC soluble to enhance binding ability. The diabody and sc(Fv)2 show high
 CC agonistic activity against anti-human Mpl antibody, such activity of which
 CC is comparable or superior to that of the natural ligand, human TPO. This
 CC is the amino acid sequence of a single chain antibody hVB22B q-wz5.
 XX
 SQ Sequence 524 AA;

Query Match 68.1%; Score 952.5; DB 9; Length 524;
 Best Local Similarity 74.8%; Pred. No. 3.8e-60;
 Matches 190; Conservative 25; Mismatches 30; Indels 9; Gaps 4;
 QY 2 AQVLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTH 59
 DB 19 SQVLVQSGPEVKKPGASVKVSCKASGYTFTNSWNVRQPGKLEWIGRIYPGDGETI 78
 QY 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDDE---PYWQGGTLVTV 116
 DB 79 YNGKFRVVTITADESTSTAYMELSLRSDDTAVVYCA---RGYDDYSFAYWGGTITV 135
 QY 117 SSGGGGGGGGGGGGGSEIVLTQSPSLPVLTPVLTQSPASISCRSSQNLVYSGNTYLSWFOQ 176
 DB 136 SSGGGGGGGGGGGGGSDIVMTQSPSLPVLTPVTPGPASISCRSSKSLHSGNTYLYWFOQ 195
 QY 177 RPSQSPRLIYKVNRSQVDPDRPSGSGGTDFTLKISRVEAEDIGVYCMQGTTHPPRT 236
 DB 196 KPGQAPRLIYRMSNLASGVDPDRPSGSGGTFTLKISRVEAEDVGYYCMQHIEY-PFT 254
 QY 237 FGQGTKEIKRGLG 250
 DB 255 FGQGTKEIKGGG 268

RESULT 10
 AEE02143
 ID AEE02143 standard; protein; 524 AA.
 AC AEE02143;
 XX

26-JAN-2006 (first entry)
 DE Anti-Mpl sc(Fv)2 antibody hVB22B q-wz5.
 XX
 KW Antibody; myeloproliferative leukemia virus oncogene; Mpl;
 KW antibody engineering; humanized antibody; platelet; thrombocytopenia;
 KW hemostatic; hematological disease; ds; coding sequence;
 KW single chain antibody.
 XX
 OS Homo sapiens.
 OS Mus musculus.
 OS Synthetic.
 PN WO2005107784-A1.
 XX
 PD 17-NOV-2005.
 XX
 PF 11-MAY-2005; 2005WO-JP008592.
 XX
 PR 11-MAY-2004; 2004JP-00141618.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX

PI Tsuchiya M, Orita T, Matsumoto Y;
 XX WPI; 2005-779430/79.
 DR N-PSDB; AEE02145.
 XX
 PT Remedies for thrombopenia as well as agents for increasing number of
 PT platelets comprises e.g. Mpl-recognizing antibodies with high agonistic
 PT activity against anti-human Mpl antibody.
 XX
 PS Example 3; SEQ ID NO 101; 71pp; Japanese.
 XX
 CC The invention relates to remedies for thrombocytopenia comprising an
 CC sc(Fv)2 (a double single chain antibody comprising 2 heavy chains and 2
 CC light chains, often referred to as a diabody) with binding activity
 CC toward TPO (thrombopoietin) receptor (Mpl), myeloproliferative leukemia
 CC virus oncogene) as active ingredient. The antibodies comprise human
 CC framework regions (FR) grafted to mouse complementarity region (CDR).
 CC Also included are remedies for thrombocytopenia containing antibodies
 CC that can bind with human Mpl and monkey Mpl, remedies for
 CC thrombocytopenia containing antibodies with agonistic activity against
 CC Mpl, remedies for thrombocytopenia containing antibodies (which have a
 CC heavy-chain variable region with CDR1, CDR2 and CDR3 that have amino acid
 CC sequences of AEE02043-AEE02045, respectively, and/or a light-chain
 CC variable region with CDR1, CDR2 and CDR3 that have amino acid sequences
 CC of AEE02046-AEE02048, or antibodies having amino acid sequences based on
 CC AEE02043-AEE02048 for modification by substitution, deletion, addition
 CC and/or insertion of some amino acids), remedies for thrombocytopenia
 CC containing antibodies that can recognize amino acid sites of human Mpl
 CC from positions 126-274, and agents for increasing platelets containing
 CC the active ingredients as already specified for the thrombocytopenia
 CC remedies. The remedies are for the treatment of thrombocytopenia. Such
 CC antibodies are extremely effective because they have high agonistic
 CC activity against anti-human Mpl antibody, and the activity is comparable
 CC or superior to that of human TPO (thrombopoietin) which is a natural
 CC ligand. The present sequence represents a humanized single chain antibody
 CC (or diabody) derived from the mouse anti-Mpl antibody (VB22B). Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 524 AA;

Query Match 68.1%; Score 952.5; DB 9; Length 524;
 Best Local Similarity 74.8%; Pred. No. 3.8e-60;
 Matches 190; Conservative 25; Mismatches 30; Indels 9; Gaps 4;
 QY 2 AQVLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTH 59
 DB 19 SQVLVQSGPEVKKPGASVKVSCKASGYTFTNSWNVRQPGKLEWIGRIYPGDGETI 78
 QY 60 YAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDDE---PYWQGGTLVTV 116
 DB 79 YNGKFRVVTITADESTSTAYMELSLRSDDTAVVYCA---RGYDDYSFAYWGGTITV 135
 QY 117 SSGGGGGGGGGGGGGSEIVLTQSPSLPVLTPVLTQSPASISCRSSQNLVYSGNTYLSWFOQ 176
 DB 136 SSGGGGGGGGGGGGGSDIVMTQSPSLPVLTPVTPGPASISCRSSKSLHSGNTYLYWFOQ 195
 QY 177 RPSQSPRLIYKVNRSQVDPDRPSGSGGTDFTLKISRVEAEDIGVYCMQGTTHPPRT 236
 DB 196 KPGQAPRLIYRMSNLASGVDPDRPSGSGGTFTLKISRVEAEDVGYYCMQHIEY-PFT 254
 QY 237 FGQGTKEIKRGLG 250
 DB 255 FGQGTKEIKGGG 268

RESULT 11
 AEC92009
 ID AEC92009 standard; protein; 247 AA.
 XX
 AC AEC92009;
 XX

01-DEC-2005 (first entry)
 Single chain variable fragment antibody BC-15 protein SEQ ID NO 4.
 cytostatic; immunostimulant; pharmaceutical; protein interaction;
 therapeutic; immune modulation; solid tumor; leukemia; colorectal tumor;
 breast tumor; uterine cervix tumor; uterine fibroids; ovary tumor;
 gynecology and obstetrics; polycystic ovary syndrome; antifertility;
 gynecological; polyp; growth disorder; prostate tumor; prostate disease;
 andrology; genitourinary disease; pituitary tumor; endocrine disease;
 adenocarcinoma; melanoma; bone tumor; musculoskeletal disease;
 multiple myeloma; hematological disease; immune disorder;
 central nervous system tumor; neurological disease; glioma;
 astroblastoma; neoplasm; single chain antibody; BC-15.
 Homo sapiens.
 Synthetic.
 Key Location/Qualifiers
 Misc-difference 247
 /note= "Encoded by GATGACAAATAA"
 WO2005091805-A2.
 06-OCT-2005.
 14-FEB-2005; 2005WO-US004612.
 13-FEB-2004; 2004US-0544807P.
 10-NOV-2004; 2004US-0626726P.
 11-FEB-2005; 2005US-00056825.
 (SCRI) SCRIPPS RES INST.
 Felding-Habermann B, Janda KD, Saven A;
 WPI; 2005-664872/68.
 N-PSDB; AEC92008.
 New scFv Bc-12 or Bc-15 antibody comprises ligand mimetic, which
 specifically binds to activated alphavbeta3 integrin receptor, useful for
 treating, e.g. neoplastic disease, solid tumor, leukemia, or benign or
 malignant breast cancer.
 Claim 3; SEQ ID NO 4; 109pp; English.
 The invention describes an antibody comprising a ligand mimetic, which
 specifically binds to an activated alphavbeta 3 integrin receptor
 differentially produced on a cell in a metastatic state compared to a
 similar, non-metastatic cell, and does not bind to a non-activated
 alphavbeta 3 integrin receptor. Also described are: a pharmaceutical
 composition comprising the antibody; treating a disease state in a mammal
 ; a cell line comprising a tumor cell variant with a metastatic homing
 propensity to a target tissue; producing an antibody phage population
 having affinity for a tumor cell target; detecting tumor cells in a
 mammal by treatment with a cancer therapeutic; inducing or enhancing an
 immune response to an antigen in a mammal; detecting an activated cell
 surface receptor on a metastatic tumor cell surface in a mammalian tissue
 sample; interfering with cells liable to undergo metastasis associated
 with a disease state; identifying cells liable to undergo metastasis
 polynucleotide comprising a nucleotide sequence that has at least 90%
 percent identity to a fully defined 802 or 752 bp sequence (SEQ ID NO. 1
 or 3) given in the specification; an isolated polypeptide comprising a
 nucleotide sequence that has at least 90% sequence identity to SEQ ID NO.
 1 or 3, or shares a biological function with Bc-12 or Bc-15; a vector
 comprising the polynucleotide; an expression vector comprising the
 polynucleotide in which the nucleotide sequence of the polynucleotide is
 operatively linked with a regulatory sequence that controls expression of
 the polynucleotide in a host cell; a host cell comprising the
 polynucleotide or progeny of the cell; and determining anti-metastatic
 activity of a test compound in a mammal. The antibody is useful for
 treating neoplastic disease, solid tumor, hematological malignancy,

CC leukemia, colorectal cancer, benign or malignant breast cancer, uterine
 CC cancer, uterine leiomyomas, ovarian cancer, endometrial cancer,
 CC polycystic ovary syndrome, endometrial polyps, prostate cancer, prostatic
 CC hypertrophy, pituitary cancer, adenomyosis, adenocarcinomas, meningioma,
 CC melanoma, bone cancer, multiple myeloma, CNS cancer, glioma, or
 CC astroblastoma. This is the amino acid sequence of single chain variable
 CC fragment antibody BC-15.
 XX
 SQ Sequence 247 AA;
 Query Match 68.1%; Score 952; DB 9; Length 247;
 Best Local Similarity 71.8%; Pred No. 1.9e-60;
 Matches 188; Conservative 29; Mismatches 27; Indels 18; Gaps 5;
 QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFNNGYGITVVRQAPGGQLEWMGWIN--NGNT 58
 DQ 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFNNGYGITVVRQAPGGQLEWMGWINPDSDGT 60
 QY 59 HYAQKFGQSRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPTWGQGTLLTVSS 118
 DQ 61 NSAQQFQGRVTMTDRTSISTAYMELSLRSDDTAVYYCARDPRGDDPTWGQGTLLTVSS 120
 QY 119 GGGSGGGSGGGSGGSEIVLTQSPSLSPVTLGQPASISCRSSQNLVYSDGNTVLSWFFQORP 178
 DQ 121 G-----GGGSGSEIVLTQSPGTLSPGGERATLSCRASQSV-----SSSYLAWTQQKP 167
 QY 179 QQSPRLIYKVSNRDSGVPDFSGSGSGTFTLKISRVEAEDIGVYYCMQ--GTHWPPRTF 237
 DQ 168 QQAPRLIYGASSRATGIPDRFSGSGSGTFTLTISRLEPEDFAVYYCOQYGS--SPRTF 225
 QY 238 GQGTVKEIKRGLGLVDYKDD 259
 DQ 226 GQGTKVDIKRGLGLVDYKDD 247
 RESULT 12
 AEA44222 standard; protein; 524 AA.
 XX
 AC AEA44222;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DE Anti-TPO single chain antibody hVB22B u2-wz5.
 XX
 KW hemostatic; single chain antibody; protein purification;
 KW antibody engineering; antibody production; thrombocytopenia; hemostatic;
 KW hematological disease; hVB22B u2-wz5.
 OS Homo sapiens.
 XX
 PN WO2005056604-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 10-DEC-2004; 2004WO-JP018506.
 XX
 PR 12-DEC-2003; 2003JP-00415746.
 PR 12-MAR-2004; 2004JP-00071763.
 PR 27-AUG-2004; 2004JP-00248323.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Tsunoda H, Nakano K, Orita T, Tsuchiya M, Hirata Y;
 XX
 DR WPI; 2005-445149/45.
 DR N-PSDB; AEA44221.
 XX
 PT Anti-Mpl antibodies including antihuman Mpl antibody and variants like
 PT genetically-engineered antihuman Mpl diabody, applicable in drug
 PT compositions for use in chemotherapy of thrombocytopenia and other
 PT diseases associated with Mpl mutation.
 XX

PS	Claim 29; SEQ ID NO 287; 106pp; Japanese.
XX	
CC	The invention describes an antibody is a single-chain polypeptide with
CC	TPO (Thrombopoietin) receptor (Mpl)-binding activity that contains the 2
CC	heavy-chain variable regions and 2 light-chain variable regions. Also
CC	described are: a chimeric antibody for binding with Mpl; an antibody for
CC	binding with soluble Mpl; antibodies for binding with human Mpl and
CC	monkey Mpl; antibodies with agonistic activity against human Mpl and
CC	monkey Mpl; antibodies with binding activity for soluble Mpl; antibodies
CC	with TPO agonistic activity; antibodies containing a heavy-chain variable
CC	region with CDR (complementary-determinant regions 1, 2 and 3);
CC	antibodies containing a light-chain variable region with CDR 1, 2 and 3;
CC	antibodies containing heavy and light-chain variable regions with CDR 1,
CC	2 and 3 selected from the already-specified groups in 18 combinations;
CC	antibodies containing a heavy-chain variable region with FR1, 2, 3 and 4;
CC	antibodies containing a light-chain variable region FR1, 2, 3 and 4;
CC	epitope-recognizing antibodies that can recognize the antibodies with the
CC	heavy and light-chain variable regions and amino acid sequences;
CC	antibodies that can recognize the an amino acid moiety in human Mpl from
CC	positions 26-274; polynucleotides encoding these antibodies;
CC	polynucleotides hybridizable with the polynucleotides that encode the
CC	antibodies under stringent conditions and the antibodies encoded by which
CC	have equivalent activity as such antibodies; vectors containing these
CC	polynucleotides; host cells carrying such polynucleotides or vectors; and
CC	drug compositions containing these antibodies. The antibodies are
CC	applicable in drug compositions for use in chemotherapy of
CC	thrombocytopenia and other diseases associated with Mpl mutation such as
CC	congenital amegakaryocytic thrombocytopenia. These antibodies can include
CC	anti-human Mpl antibody and variants like genetically-engineered anti-human
CC	Mpl diabody and (humanized) anti-human Mpl sc(Fv)2. Such antibodies are
CC	soluble to enhance binding ability. The diabody and sc(Fv)2 show high
CC	agonistic activity against anti-human Mpl antibody, such activity of which
CC	is comparable or superior to that of the natural ligand, human TPO. This
CC	is the amino acid sequence of a single chain antibody hVB22B u2-wz5.
XX	
SQ	Sequence 524 AA;
	Query Match. 67.9%; Score 949.5; DB 9; Length 524;
	Best Local Similarity 74.4%; Pred. No. 6.2e-60;
	Matches 189; Conservative 26; Mismatches 30; Indels 9; Gaps 4;
QY	2 AQVLVQSAGAEVKPGASVKSVCKASGYTFNNGITWVRQAPGQGLEWMGWMI--NNGNTH 59
DB	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY	19 SQQLVQSGLPEVKKPGASVKSVCKASGYTFTNSWNWVRQRPGKLEIGRIRYPGDGETI 78
QY	60 YAKFQGRVTWTDTSTNSTAYMEILRSLSDDTAVYYCARDPRGDE---PYWGQGLTVTV 116
DB	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY	79 YNGKFRVRVTITADESTAYMQJSSLRSEDTAVYYCA---RGYDDYSFAYWGQGTIVTV 135
QY	117 SSGGGSGGGSGGGGGGSEIVLTQPSLSPVTLGPASPISCRSSONLVYSDGNITSLVSFQQ 176
DB	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY	136 SSGGGSGGGSGGGGGSDIVMTQPSLSPVTPGPSAISCRSSKLHNSGNTLYLWFQ 195
QY	177 RFGOSPRRLIYKSNRDSGVDPDRFSGGSGDFTFLKISRVEADIDGVYCMQGTHWPPT 236
DB	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY	196 KPGOSPQLLIYMSNLASGVDPDRFSGGSGDFTFLKISRVEADIDGVYCMQHIEY-PPT 254
QY	237 FGQGTKVEIKRGLG 250
DB	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY	255 FGQGTKLEIKRGGG 268
RESULT 13	
AEE02142	
ID	AEE02142 standard; protein; 524 AA.
XX	
AC	AEE02142;
XX	
DT	26-JAN-2006 (first entry)
XX	
DE	Anti-Mpl sc(Fv)2 antibody hVB22B u2-wz4 .
XX	
KW	Antibody; myeloproliferative leukemia virus oncogene; Mpl;

KW antibody engineering; humanized antibody; platelet; thrombocytopenia;
KW hemostatic; hematological disease; ds; coding sequence;
KW single chain antibody.

OS Homo sapiens.
OS Mus musculus.
OS Synthetic.

XX WO2005107784-A1.
XX
PD 17-NOV-2005.
XX
XX 11-MAY-2005; 2005WO-JP008592.
FF
XX
PR 11-MAY-2004; 2004JP-0014618.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
FI Tsuchiya M, Orita T, Matsumoto Y;
XX
DR WPI; 2005-779430/79.
DR N-PSTB; AEE02144.
XX

Remedies for thrombopenia as well as agents for increasing number of platelets comprises e.g. Mpl-recognizing antibodies with high agonistic activity against anti-human Mpl antibody.

PS Example 3; SEQ ID NO 100; 71pp; Japanese.

The invention relates to remedies for thrombocytopenia comprising an sc(Fv)2 (a double single chain antibody comprising 2 heavy chains and 2 light chains, often referred to as a diabody) with binding activity toward TPO (thrombopoietin) receptor (Mpl), myeloproliferative leukemia virus oncogene) as active ingredient. The antibodies comprise human framework regions (FR) grafted to mouse complementarity region (CDR). Also included are remedies for thrombocytopenia containing antibodies that can bind with human Mpl and monkey Mpl, remedies for thrombocytopenia containing antibodies with agonistic activity against Mpl, remedies for thrombocytopenia containing antibodies (which have a heavy-chain variable region with CDR1, CDR2 and CDR3 that have amino acid sequences of AEE02043-AEE02045, respectively, and/or a light-chain variable region with CDR1, CDR2 and CDR3 that have amino acid sequences of AEE02046-AEE02048, or antibodies having amino acid sequences based on AEE02043-AEE02048 for modification by substitution, deletion, addition and/or insertion of some amino acids), remedies for thrombocytopenia containing antibodies that can recognize amino acid sites of human Mpl from positions 126-274, and agents for increasing platelets containing the active ingredients as already specified for the thrombocytopenia remedies. The remedies are for the treatment of thrombocytopenia. Such antibodies are extremely effective because they have high agonistic activity against anti-human Mpl antibody, and the activity is comparable or superior to that of human TPO (thrombopoietin) which is a natural ligand. The present sequence represents a humanized single chain antibody (or diabody) derived from the mouse anti-Mpl antibody (VB22B). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 524 AA;

SQ Query Match 67.9%; Score 949.5; DB 9; Length 524;
Best Local Similarity 74.4%; Pred. No. 6.2e-60;
Matches 189; Conservative 26; Mismatches 30; Indels 9; Gaps 4;

Qy 2 AQVOLVGSGAVKFKPGASVKVCSCASGYTFNSYGITWRAPCGGLEWMGWI--NNGNTH 59
Dd :|||||:
19 SQVOLVQSPEVKPKPGASVKVCSCASGYTFTNSMNWNVRPFGKGLEWIGRIYPGDGETI 78
Qy 60 YAQKFQGRTVTMTDTSTATYWELRSLSRDDTAIVYYCARDPRGDDDE---PYWGQGTLTVV 116
Dd 79 YNGKFRVRTVIADESTISTAMQLSLSESDATAVYYCA---RGYDDSFAYWGQGTTVTIV 135
Qy 117 SSSGGGSGGGGGGGGEISEVLTVGPGLSPLVTLTGQPASI:SCRSSQNLVDSGNTYLISWFQQ 176

PT Remedies for thrombopenia as well as agents for increasing number of
PT platelets comprises e.g. Mpl-recognizing antibodies with high agonistic
PT activity against antihuman Mpl antibody.
XX
XX
XX Example 3; SEQ ID NO 61; 71pp; Japanese.
XX
CC The invention relates to remedies for thrombocytopenia comprising an
CC sc(Fv)2 (a double single chain antibody comprising 2 heavy chains and 2
CC light chains, often referred to as a diabody) with binding activity
CC toward TPO (thrombopoietin) receptor (Mpl, myeloproliferative leukemia
CC virus oncogene) as active ingredient. The antibodies comprise human
CC framework regions (FR) grafted to mouse complementarity region (CDR).
CC Also included are remedies for thrombocytopenia containing antibodies
CC that can bind with human Mpl and monkey Mpl, remedies for
CC thrombocytopenia containing antibodies with agonistic activity against
CC Mpl, remedies for thrombocytopenia containing antibodies (which have a
CC heavy-chain variable region with CDR1, CDR2 and CDR3 that have amino acid
CC sequences of AEE02043-AEE02045, respectively, and/or a light-chain
CC variable region with CDR1, CDR2 and CDR3 that have amino acid sequences
CC of AEE02046-AEE02048, or antibodies having amino acid sequences based on
CC AEE02043-AEE02048 for modification by substitution, deletion, addition
CC and/or insertion of some amino acids), remedies for thrombocytopenia
CC containing antibodies that can recognize amino acid sites of human Mpl
CC from positions 126-274, and agents for increasing platelets containing
CC the active ingredients as already specified for the thrombocytopenia
CC remedies. The remedies are for the treatment of thrombocytopenia. Such
CC antibodies are extremely effective because they have high agonistic
CC activity against antihuman Mpl antibody, and the activity is comparable
CC or superior to that of human TPO (thrombopoietin) which is a natural
CC ligand. The present sequence represents a humanized single chain antibody
CC (or diabody) derived from the mouse anti-Mpl antibody (VB22B). Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 524 AA;

Query Match	67.8%;	Score	947.5;	DB	9;	Length	524;
Best Local Similarity	74.8%;	Pred.	No. 8.6e-60;				
Matches	190;	Conservative	24;	Mismatches	31;	Indels	9;
						Gaps	4;
QY	2	AQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTH	59				
Db	19	SQVQLVQSGPEVKKPGASVKVSCKASGYTFNWMNVRQPGKLEWMGRIYPGDGETI	78				
QY	60	YAKFQGRVTMTTDTSTAYMELSLRSDDTAYYICARDPRGDDE---PYWGQGLTVTV	116				
Db	79	YNGKFRVVTITADESTAYMELSLRSDBTAYYICA---RGYDDYSFAYWGQGLTVTV	135				
QY	117	SSGGSGSGSGSGSGSEIYLTSQSLPLVTLGQPASISCRSSONLVYSDGNTYLSWFOQ	176				
Db	136	SSGGSGSGSGSGSGGGSDIVMTQALSLVTPPGPAPASISCRSSKSLHSNGNTLYWFOQ	195				
QY	177	RPQSPRLIYKVSNRDSGVDPDRFSGSGSGDTFTLKISRVEAEDIGVYCMQGTWPPPT	236				
Db	196	KPGQSPQLLIYRHSNLSAGVDPDRFSGSGSGDTFTLKISRVEAEDVGVYCMQHIEY-PPT	254				
QY	237	FGQGTKVEIKRGLG	250				
Db	255	FGQGTKLEIKGGGG	268				

Search completed: August 9, 2006, 13:09:33
Job time : 198 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:14:57 ; Search time 50 Seconds
(without alignments)
456.910 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC_Celerra_SID33/ptodata/2/iaa/5_COMB.pep.*
- 2: /EMC_Celerra_SID33/ptodata/2/iaa/6_COMB.pep.*
- 3: /EMC_Celerra_SID33/ptodata/2/iaa/7_COMB.pep.*
- 4: /EMC_Celerra_SID33/ptodata/2/iaa/H_COMB.pep.*
- 5: /EMC_Celerra_SID33/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /EMC_Celerra_SID33/ptodata/2/iaa/RE_COMB.pep.*
- 7: /EMC_Celerra_SID33/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	914.5	65.4	535	2	US-08-983-035A-38
2	914.5	65.4	535	2	US-09-968-851A-38
3	909.5	65.1	243	2	US-09-297-181-2
4	905	64.7	267	2	US-09-419-788-30
5	883	63.2	353	2	US-09-203-958A-4
6	881.5	63.1	365	2	US-08-875-811-53
7	881.5	63.1	366	2	US-08-875-811-55
8	870	62.2	257	2	US-09-419-788-113
9	866	61.9	281	2	US-09-025-769B-178
10	866	61.9	281	2	US-09-490-070A-178
11	866	61.9	281	2	US-09-490-153-178
12	866	61.9	281	2	US-09-490-324-178
13	856	61.2	249	3	US-09-661-992B-84
14	850	60.8	249	2	US-09-726-219A-190
15	850	60.8	249	2	US-09-196-522-190
16	848	60.7	269	2	US-08-646-265A-109
17	828	59.2	247	2	US-09-227-693-34
18	828	59.2	248	1	US-08-331-397B-34
19	828	59.2	248	1	US-08-331-397B-34
20	828	59.2	248	1	US-08-759-804A-34
21	828	59.2	260	1	US-08-447-402-1
22	826	59.1	269	2	US-09-070-408-132
23	823	58.9	263	1	US-08-752-844-66
24	823	58.9	263	2	US-09-293-533-66
25	821	58.7	673	2	US-09-423-439-32
26	816	58.4	252	1	US-08-133-804-4

27	816	58.4	252	1	US-08-461-838-4	Sequence 4, Appli
28	816	58.4	252	1	US-08-461-838-4	Sequence 4, Appli
29	816	58.4	288	2	US-09-423-439-38	Sequence 38, Appl
30	808	57.8	246	1	US-08-257-341-7	Sequence 7, Appli
31	808	57.8	367	1	US-08-257-341-5	Sequence 5, Appli
32	807	57.7	289	2	US-09-184-658-63	Sequence 63, Appl
33	807	57.7	289	2	US-09-504-262D-63	Sequence 63, Appl
34	802	57.4	267	2	US-09-485-737B-2	Sequence 2, Appli
35	802	57.4	267	2	US-10-071-485-2	Sequence 2, Appli
36	794	56.8	711	2	US-09-485-737B-90	Sequence 90, Appl
37	794	56.8	711	2	US-10-071-485-90	Sequence 90, Appl
38	793	56.7	541	2	US-09-485-737B-85	Sequence 85, Appl
39	793	56.7	541	2	US-10-071-485-85	Sequence 85, Appl
40	790	56.5	240	1	US-08-488-113B-148	Sequence 148, App
41	790	56.5	240	1	US-08-477-484B-148	Sequence 148, App
42	790	56.5	240	1	US-08-646-360-148	Sequence 148, App
43	790	56.5	240	2	US-08-839-765-148	Sequence 148, App
44	790	56.5	240	2	US-09-136-389-148	Sequence 148, App
45	790	56.5	240	2	US-09-610-838-148	Sequence 148, App

ALIGNMENTS

RESULT 1

US-08-983-035A-38
; Sequence 38, Application US/08983035A
; Patent No. 6326464
; GENERAL INFORMATION:
; APPLICANT: CONSELLER, EMMANUEL
; TITLE OF INVENTION: P53 PROTEIN VARIANTS AND THERAPEUTICAL
; USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983,035A
; FILING DATE: 20-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/01111
; FILING DATE: 17-JUL-1996
; APPLICATION NUMBER: FR 95/08729
; FILING DATE: 19-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Strauss, William L.
; REGISTRATION NUMBER: 47,114
; REFERENCE/DOCKET NUMBER: 03804.0142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-08-983-035A-38

Query Match 65.4%; Score 914.5; DB 2; Length 535;

[illegible]

RESULT 2
 US-09-968-851A-38
 ; Sequence 38, Application US/09968851A
 ; Patent No. 6933373
 ; GENERAL INFORMATION:
 ; APPLICANT: CONSEILLER, EMMANUEL
 ; APPLICANT: BRACCO, LAURENT
 ; TITLE OF INVENTION: P53 PROTEIN VARIANTS AND THERAPEUTICAL
 ; USES THEREOF
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FINNEGAN, HENDERSON, PARABOW, GARRETT &
 ; DUNNER, LLP
 ; STREET: 1300 I Street, NW
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/968,851A
 ; FILING DATE: 03-Oct-2001
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/983,035
 ; FILING DATE: 20-FEB-1998
 ; APPLICATION NUMBER: PCT/FR96/01111
 ; FILING DATE: 17-JUL-1996
 ; APPLICATION NUMBER: FR 95/08729
 ; FILING DATE: 19-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Strauss, William L.
 ; REGISTRATION NUMBER: 47,114
 ; REFERENCE/DOCKET NUMBER: 03804.0142-01000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-408-4000
 ; TELEFAX: 202-408-4400
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 535 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
 US-09-968-851A-38

	Query Match	65.4%; Score 914.5; DB 2; Length 535;
	Best Local Similarity	69.9%; Pred. No. 3.7e-67;
	Matches	174; Conservative 34; Mismatches 36; Indels 5; Gaps 3;
Qy	1 MAOVLVOSGAELVKGASVKVSKAGSYTFSNYGITWVRQAPGGGLEWMGWIN--NGNT	58
Dd	1 MAOVQLQESGAELVSGASVKLSCTASGNFIKDYIMHWVKQRPEQGLEWIGWIDPENGDT	60
Qy	59 HYAQKFQGRVTMTTSTSTAYMELSLRSDDTAIVYICARDPRGDDEPYWGOGTLTVSS	118
Dd	61 EYAPKFGKATMTADTSNSTAYILQLSSLASEDAVYYC--NFYGDALDLYWGOGTTTVSS	118
Qy	119 GGCGSGGGGGGGGSRIIVLTQSPLSLPVLTIGQPASISCRSSQNLYVSDGNTYLSWFQORP	178
Dd	119 GGCGSGGGGGGGGSVDLMTQTPTLLSVTIQGPASISCKSSQSLDSDGKTYLNWLLQRP	178
Qy	179 GOSPRLIYKVNSRDGVDPFSGSGGTDFTLKISRVEAEDIGVYCMQGTHWPRTFG	238
Dd	179 GOSPRLIYVLVSKLDGVPDRFTGSGSGTDFTLKINRVEAEDLGVIYCWQGTH-SPLTFG	237
Qy	239 QGTKVEIK 247	
Dd	238 AGTKLEIK 246	
RESULT 3		
US-09-297-181-2		
; Sequence 2, Application US/09297181		
; Patent No. 6852509		
; GENERAL INFORMATION:		
; APPLICANT: Bracco, Laurent		
; TITLE OF INVENTION: ANTI-P53 SINGLE-CHAIN ANTIBODY FRAGMENTS AND THEIR USES		
; FILE REFERENCE: ST96030-US		
; CURRENT APPLICATION NUMBER: US/09/297,181		
; CURRENT FILING DATE: 1999-04-26		
; EARLIER APPLICATION NUMBER: PCT/FR97/01921		
; EARLIER FILING DATE: 1997-10-27		
; EARLIER APPLICATION NUMBER: FR96/13176		
; EARLIER FILING DATE: 1996-10-29		
; NUMBER OF SEQ ID NOS: 4		
; SOFTWARE: Patent In Ver. 2.0		
; SEQ ID NO 2		
; LENGTH: 243		
; TYPE: PRT		
; ORGANISM: Mus musculus		
US-09-297-181-2		
Query Match		
Best Local Similarity 65.1%; Score 909.5; DB 2; Length 243;		
Matches 173; Conservative 33; Mismatches 35; Indels 5; Gaps 3;		
Qy	3 QVOLVOSGAELVKGASVKVSKAGSYTFSNYGITWVRQAPGGGLEWMGWIN--NGNTHY	60
Dd	1 QVOLQSGAELVSGASVKLSCTASGNFIKDYIMHWVKQRPEQGLEWIGWIDPENGDT	60
Qy	61 AQEFGQGRVTMTTSTSTAYMELSLRSDDTAIVYICARDPRGDDEPYWGOGTLTVSSGG	120
Dd	61 APFQGGKATMTADTSNTAYILQLSSLASEDAVYYC--NFYGDALDLYWGOGTTTVSSGG	118
Qy	121 GGSGGGGGGGGGSEIIVLTQSPLSLPVLTIGQPASISCRSSQNLYVSDGNTYLSWFQORPGQ	180
Dd	119 GGSGGGGGGGGGSDVLMTQTPTLLSVTIQGPASISCKSSQSLDSDGKTYLNWLLQRPQ	178
Qy	181 SPRLIYKVNSRDGVDPFSGSGGTDFTLKISRVEAEDIGVYCMQGTHWPRTFGQ	240
Dd	179 SPKRLLYLVSKLDGVPDRFTGSGSGTDFTLKINRVEAEDLGVIYCWQGTH-SPLTFEG	237
Qy	241 TKVEIK 246	
Dd	238 TKLEIK 243	

RESULT 4

US-09-419-788-30
; Sequence 30, Application US/09419788
; Patent No. 6825325
; GENERAL INFORMATION:
; APPLICANT: FISCHER, Rainer
; APPLICANT: SCHILLBERG, Stefan
; APPLICANT: NAHRING, Jorg
; APPLICANT: SACK, Markus
; APPLICANT: MONECKE, Michael
; APPLICANT: LIAO, Yu-Gai
; APPLICANT: SPIEGEL, Holger
; APPLICANT: ZIMMERMAN, Sabine
; APPLICANT: EMANS, Neil

; TITLE OF INVENTION: Molecular Pathogenicide Mediated Plant Disease

; FILE REFERENCE: 0147-0189P

; CURRENT APPLICATION NUMBER: US/09/419,788

; EARLIER FILING DATE: 1999-10-18

; EARLIER FILING DATE: 1998-10-16

; EARLIER FILING DATE: 1998-10-16

; NUMBER OF SEQ ID NOS: 163

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 30

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic, no

; OTHER INFORMATION: natural origin

US-09-419-788-30

Query Match 64.7%; Score 905; DB 2; Length 267;

Best Local Similarity 68.5%; Pred. No. 1e-66;

Matches 176; Conservative 34; Mismatches 43; Indels 4; Gaps 3;

3 QVQLVQSGAEVKPGASVKVSCKASGYTFNIGITWVRQAPGQGLEWMGWIN--NGNTHY 60

1 EVKLQSGAEVLVPGASVKLSCKASDYTFGYNMWVKQSHGKSLWGINPNPYGTSY 60

61 AQFQGRVTMTDTSTSTAYMELRLSRDSDTAVYYCARDPRGDD--EPYWGQGLTVTVSSG 119

61 NQFQKQKATLTVDKSSSTAYMQLNSLTSDSAVYCAVGGYVDWFAVWGQGLTVTVSSG 120

120 GGGSGGGSGGGGSEIVLTQSPVLTLPVLTQSPASISCRSSQSLVHSGNTYHLWYLNQPG 179

121 GGGSGGGSGGGGSDILLTQSPVLTLPVLTQSPASISCRSSQSLVHSGNTYHLWYLNQPG 180

180 QSPRLIYKVSNRDGVDPDRFGSGGSGTDTLTLSRVEAEDIGVYVCWGQTHWPPRTFGQ 239

181 QSPRLIYKVSNRDGVDPDRFGSGGSGTDTLTLSRVEAEDIGVYVCWGQTHWPPRTFGQ 239

240 GTKVEIKRGLGLVDYK 256

240 GTKLEKRAVDAAEQK 256

RESULT 5

US-09-203-958A-4
; Sequence 4, Application US/09203958A
; Patent No. 6682928
; GENERAL INFORMATION:
; APPLICANT: KELLER, Tibor
; APPLICANT: GOLDSTEIN, Joel
; APPLICANT: GRAZIANO, Robert
; APPLICANT: DBO, Yashwant M.

; TITLE OF INVENTION: CELLS EXPRESSING ANTI-PC RECEPTOR

; FILE REFERENCE: MXI-099CPA

; CURRENT APPLICATION NUMBER: US/09/203,958A

; CURRENT FILING DATE: 1998-12-02

; PRIOR APPLICATION NUMBER: 60/067232

; PRIOR FILING DATE: 1997-12-02

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic construct

US-09-203-958A-4

Query Match 63.2%; Score 883; DB 2; Length 353;

Best Local Similarity 66.1%; Pred. No. 8.9e-65;

Matches 168; Conservative 39; Mismatches 35; Indels 12; Gaps 4;

2 AQVQLVQSGAEVKPGASVKVSCKASGYTFNIGITWVRQAPGQGLEWMGWIN--GN 57

37 SEIQQTGPGLVPGASVKLSCKASGYTFYIIFWVKQSHGKSLWGTG--NNNPYGS 94

58 THYAKQFQGRVTMTDTSTSTAYMELRLSRDSDTAVYYCAR-----DPRGDDPEYWGQGT 112

95 TSYNLKFKGKATLTVDKSSSTAYMQLNSLTSDSAVYCVRGVYVYSSVYEAFFYWGQGT 154

113 LVTSSGGSGGGGSGGSEIVLTQSPVLTLPVLTQSPASISCRSSQSLVHSGNTYHL 172

155 LVTSSGGSGGGGSGGSDVMTQTPVLTLSITIGTPASISCRSSQSLVHSGNTYHL 214

173 WFOQRPQOSPRRLIYKVSNRDGVDPDRFGSGGSGTDTLTLSRVEAEDIGVYVCWGQTHW 232

215 WLLQRPQOSPRRLIYKVSNRDGVDPDRFGSGGSGTDTLTLSRVEAEDIGVYVCWGQTHW 274

233 PPTFGGQTKVEIK 246

275 -PPTFGGQTKVEIK 287

RESULT 6

US-08-875-811-53

; Sequence 53, Application US/08875811

; Patent No. 6045793

; GENERAL INFORMATION:

; APPLICANT: Rybak, Susanna M.

; APPLICANT: Newton, Dianne L.

; APPLICANT: Boque, Lluís

; APPLICANT: Wlodawer, Alexander

; TITLE OF INVENTION: Recombinant Ribonuclease Proteins

; NUMBER OF SEQUENCES: 64

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/875,811

; FILING DATE: 19-FEB-1998

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US97/02588

; FILING DATE: 19-FEB-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/011,800

; FILING DATE: 21-FEB-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Paris, Susan K.

; REGISTRATION NUMBER: 41,739

; REFERENCE/DOCKET NUMBER: 015280-244100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-875-811-53

Query Match 63.1%; Score 881.5; DB 2; Length 365;
Best Local Similarity 67.3%; Pred. No. 1.2e-64;
Matches 167; Conservative 37; Mismatches 37; Indels 7; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWVQAPGQGLEWGMWIN--NGNTHY 60
DB 119 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWVQAPGQGLEWGMWIN--NGNTHY 178
QY 61 AQKFGQRTVMTDTSTAYMELRSLRSDDTAVYYCARDP--RGDDEPYWGQGLTVTVSSG 119
DB 179 ADDFKGRFAFSLTSAAYLIQINLNKNEDEATYFCARFAIKGD---YWGQGLTVTVSSG 235
QY 120 GGGSGGGSGGGSGGGSEIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSWFOQRP 179
DB 236 GGGSGGGSGGGSGGGSDIVLTQSPFNPVTLTGTSASISCRSTKSLHNSGITYLYWYLRKPG 295
QY 180 QSPRLIYKVNRSQVDPDRSGSGSGTDFTLKISRVEAEIDGVYCMQGTHTWPPRTFGQ 239
DB 296 QSPQLLIYQMSNLASGVDPDRFSSSGSGTDFTLRISRVEAEIDGVYCAQNLLE--IPRTFGG 354
QY 240 GTKVEIKR 247
DB 355 GTKLEIKR 362

RESULT 7
US-08-875-811-55
; Sequence 55, Application US/08875811
; Patent No. 6045793
; GENERAL INFORMATION:
; APPLICANT: Rybak, Susanna M.
; APPLICANT: Newton, Dianne L.
; APPLICANT: Boque, Lluis
; APPLICANT: Wlodawer, Alexander
; TITLE OF INVENTION: Recombinant Ribonuclease Proteins
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,811
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/02588
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/011,800
; FILING DATE: 21-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Paris, Susan K.
; REGISTRATION NUMBER: 41,739

; REFERENCE/DOCKET NUMBER: 015280-244100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-875-811-55

Query Match 63.1%; Score 881.5; DB 2; Length 366;
Best Local Similarity 67.3%; Pred. No. 1.2e-64;
Matches 167; Conservative 37; Mismatches 37; Indels 7; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWVQAPGQGLEWGMWIN--NGNTHY 60
DB 2 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWVQAPGQGLEWGMWIN--NGNTHY 61
QY 61 AQKFGQRTVMTDTSTAYMELRSLRSDDTAVYYCARDP--RGDDEPYWGQGLTVTVSSG 119
DB 62 ADDFKGRFAFSLTSAAYLIQINLNKNEDEATYFCARFAIKGD---YWGQGLTVTVSSG 118
QY 120 GGGSGGGSGGGSGGGSEIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSWFOQRP 179
DB 119 GGGSGGGSGGGSGGGSDIVLTQSPFNPVTLTGTSASISCRSTKSLHNSGITYLYWYLRKPG 178
QY 180 QSPRLIYKVNRSQVDPDRSGSGSGTDFTLKISRVEAEIDGVYCMQGTHTWPPRTFGQ 239
DB 179 QSPQLLIYQMSNLASGVDPDRFSSSGSGTDFTLRISRVEAEIDGVYCAQNLLE--IPRTFGG 237
QY 240 GTKVEIKR 247
DB 238 GTKLEIKR 245

RESULT 8
US-09-419-788-113
; Sequence 113, Application US/09419788
; Patent No. 6825325
; GENERAL INFORMATION:
; APPLICANT: FISCHER, Rainer
; APPLICANT: SCHILLBERG, Stefan
; APPLICANT: NAHRING, Jorg
; APPLICANT: SACK, Markus
; APPLICANT: MONECKE, Michael
; APPLICANT: LIAO, Yu-Cal
; APPLICANT: SPIEGEL, Holger
; APPLICANT: ZIMMERMAN, Sabine
; APPLICANT: EMANS, Neil
; TITLE OF INVENTION: Molecular Pathogenicide Mediated Plant Disease
; TITLE OF INVENTION: Resistance
; FILE REFERENCE: 0147-0189P
; CURRENT APPLICATION NUMBER: US/09/419,788
; CURRENT FILING DATE: 1999-10-18
; EARLIER APPLICATION NUMBER: 98 11 9630.6 EP
; EARLIER FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: 66/BOM/1998 INDIA
; EARLIER FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic, no
; OTHER INFORMATION: natural origin
US-09-419-788-113
Query Match 62.2%; Score 870; DB 2; Length 257;
Best Local Similarity 66.8%; Pred. No. 7.3e-64;

Matches 171; Conservative 33; Mismatches 36; Indels 16; Gaps 5;

QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNT 58
DB 1 MAEVQLQDSGAEVKKPGASVKVSCKASGYTFTNVMHWVKQTTPGQGLEWIGAIYPRNGDT 60
QY 59 HYAQKFGQRTMTTDTSTSTAYMELRLSLRSDTAVYYCARDPRGDDPVPWGQGLTLVTYS 118
DB 61 SYNQKFGKATLTADKSSSTAYMQLSSLTSDSAVYYCAR-----PDVWAGTLLTYS 114
QY 119 GGG---SGS-----GGSGGGGSGSEIVLTQSPLSLPTVLGQPASISCRSSQNLVYSDGNTYL 171
DB 115 GAGTSGSGKPGPEGSTKGPADVLMTQAPLTVTIGQPASISCKSSQSLDGDGKTYL 174
QY 172 SWFOQRPGSPRLIYKVNRLSDGVPDRFSGSGGTDTFTLKISRVEAEDIGVYYCMQGT 231
DB 175 NWLLQRPGSPRLIYLVSKLSDGVPDRFTGSGGTDTFTLKISRVEAEDLGYYTCMQGT 234
QY 232 WPPRTFGQTKVEIKR 247
DB 235 F-PHTFGGKLEIKR 249

RESULT 9
US-09-025-769B-178
; Sequence 178, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-178

Query Match 61.9%; Score 866; DB 2; Length 281;
Best Local Similarity 66.4%; Pred. No. 1.7e-63;
Matches 170; Conservative 36; Mismatches 36; Indels 14; Gaps 5;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNT 60
DB 26 EVQLVESGGGLVQPGGSLRLSCAASGFTFSSVAMSWVRQAPGKLEWVSAISGSGSTYY 85
QY 61 AQKFGQRTMTTDTSTSTAYMELRLSLRSDTAVYYCARDPRGDD---EPYWGQGLTLTV 116
DB 86 ADSVGRFTISRDNSSKNTLYLQWNSLRADTAVYYCAR--WGSDGFYAMDYWGQGLTLTV 143
QY 117 S-----SGGGSGGGSGGGSGSEIVLTQSPLSLPTVLGQPASISCRSSQNLVYSDGNTYL 171
DB 144 SAAGSGSGGGSGGGSGGGSGGSDIVMTQSPLSLPTVTPGEPASISCRSSQSLHNSGNYL 203
QY 172 SWFOQRPGSPRLIYKVNRLSDGVPDRFSGSGGTDTFTLKISRVEAEDIGVYYCMQGT 231
DB 204 DWLQKPGSPQLLYLGNRASGVPDRFSGSGGTDTFTLKISRVEADVGYYCQHYT 263
QY 232 WPPRTFGQTKVEIKR 247
DB 264 TPP-TFGQTKVEIKR 278

RESULT 10
US-09-490-070A-178
; Sequence 178, Application US/09490070A
; Patent No. 6696248
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Colin G. Sandercock, Esq. c/o Heller Ehrman
; STREET: 1666 K Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,070A
; FILING DATE: 24-Jan-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Colin G. Sandercock, Esq.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 37629-0005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 912-2000
; TELEFAX: (202) 912-2020
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-09-490-070A-178

Query Match 61.9%; Score 866; DB 2; Length 281;
Best Local Similarity 66.4%; Pred. No. 1.7e-63;
Matches 170; Conservative 36; Mismatches 36; Indels 14; Gaps 5;

Qy		3	QVQLVQSGLAEVKPGASVKVSCKASGTYFSSYNGITWVRQAPEQGLEWNGWI--NNGNTHY	60
Db		26	EVQLVESGGGLVQPGGSLRLSCAASGTFSSYAMSVWRQAPKGLIEWSAISGSGSTYY	85
Qy		61	AQKFQGRVTMTDTSTAYMELSRSDRTAVYYCARDPRGDD---EPYWGQGTFLTVV	116
Db		86	ADSVKGFTISRDNKNTLYIQMNSLRAEDTAVYYCAR--WGGDGFYAMDYWGQGTFLTVV	143
Qy		117	S-----SGGGSGGGSGGGGSEIVLTQSPILPTLTLPQTLPASISCRSSONLVSYDGNTYL	171
Db		144	SSAGGSGGGSGGGSGGGGSDVMVTQSPISLPLPTPGEPAISICRSSQSLHSHSNLYL	203
Qy		172	SWFOQRPGQPRRLIYKVSNRDSGVDPDFSGSGSCTDFTLKLSRYEADI GVVYCMQGTH	231
Db		204	DWTLOKFGQPPQLLIYIGSNRASGVDPDRFSGSGSCTDFTLKLSRYEADVGVYCOQHYY	263
Qy		232	WPRTFGQGTKVEIKR	247
Db		264	TTP--TFGGTKVEIKR	278

```

RESULT 11
US-09-490-153-178
; Sequence 178, Application US/09490153
; Patent No. 6706484
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; Pack, Peter
; Ilag, Vic
; Ge, Liming
; Moroney, Simon
; Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/490,153
; FILING DATE: 24-Jan-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-09-490-153-178

```

Query Match 61.9%; Score 866; DB 2; Length 281;
Best Local Similarity 66.4%; Pred. NO. 1.7e-63;
Matches 170; Conservative 36; Mismatches 36; Indels 14; Gaps 5;

Qy	3	QVQLVQSGAEVKKPGASLVTKVSKAGSYTFSSYGITWRAQFGQGLEWMGWI--NNGNTHY	60
Db	26	EVQLVESGGGLVQPGRSLISCAASGFTFESSYAMSWVRQAQPKGLEWMSAISGSGGSTYY	85
Qy	61	AQKFGQGVTTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD---EPYWGQGTTLTVT	116
Db	86	ADSVKGRTISRDNASKNTLVLQNSLSRAEDTAVYYCAR--WGGDGFYAMDYWGQGTTLTVT	143
Qy	117	S-----SGGGSGGGSGGGGSEIVLTQSPLSPVTLGQPAASICRSSQNLYSDGNITYL	171
Db	144	SSAGGSGGGSGGGSGGGSGGSDIVMTQTSPLSPVTPGPAPASICRSSQSLILHSNGNYL	203
Qy	172	SWFOORPQSQPRRLIKYVSNRDSVPDRFSGSGSSTDFTLKISRVEADI GVVYCMQGTH	231
Db	204	DWYLQKPGQSPOLLIIYLGSNRASGVDPDRFSGSGSSTDFTLKISRVEADVGVYCOQHVT	263
Qy	232	WPPTFGQGTKVEIKR	247
Db	264	TTP--TEGGQTKVEIKR	278

RESULT 12
 US-09-490-324-178
 ; Sequence 178, Application US/09490324
 ; Patent No. 6828422
 ; GENERAL INFORMATION:
 ; APPLICANT: Knappik, Achim
 ; Pack, Peter
 ; Ilag, Vic
 ; Ge, Liming
 ; Moroney, Simon
 ; Plueckthun, Andreas
 ; TITLE OF INVENTION: Protein/(poly)peptide libraries
 ; NUMBER OF SEQUENCES: 373
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
 ; STREET: 1251 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10021
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/490,324
 ; FILING DATE: 24-Jan-2000
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/025,769
 ; FILING DATE: 18-FEB-1998
 ; APPLICATION NUMBER: EP 95 11 3021.0
 ; FILING DATE: 18-AUG-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: James F. Haley, Jr., Esq.
 ; REGISTRATION NUMBER: 27,794
 ; REFERENCE/DOCKET NUMBER: MORPHO/5
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212)596-9000
 ; TELEFAX: (212)596-9090
 ; INFORMATION FOR SEQ ID NO: 178:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 281 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 178:
 US-09-490-324-178

Query Match 61.9%; Score 866; DB 2; Length 281;
Best Local Similarity 66.4%; Pred. No. 1.7e-63;

Matches	170;	Conservative	36;	Mismatches	36;	Indels	14;	Gaps	5
Qy	3	QVQLVDSGAEVKKPGASVKVSKASGYTFSNYGITWVROAPGQGLEWMGWI	--NNGNTHY	60					
Db	26	EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKLEWVSAISGSGSTYY	85						
Qy	61	AQKPGQGRVTMTDTSTSTAYMELRSLRSDDTAVYVYCARDPRGDD	----EPYWGQGITLVTV	116					
Db	86	ADSVKGRFTISRDNKNTLYLQWNSLRRAEDTAVVYICAR	--WGGDGFYAMDYWGQGITLVTV	143					
Qy	117	S-----SGGGSGGGSGGGSGGSEIVLTQSPSLSPVTLTGQPAISCRSSQNLVYSDGNLYL	171						
Db	144	SSAGSGGGSGGGSGGGSGGGSDIVMQSPSLSPVTPGEPASISCRSSQSLHNSGNLYL	203						
Qy	172	SWFQQRPGQSPLRLIYKVSNRDGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYVCMQGT	231						
Db	204	DWYLQKPGQSPQLLIYLSNRASGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYVCMQGT	263						
Qy	232	WPRTPGQTKVEIKR	247						
Db	264	TPP-TFQGGTKVEIKR	278						
RESULT 13									
US-09-661-992B-84									
; Sequence 84, Application US/09661992B									
; Patent No. 7033590									
; GENERAL INFORMATION:									
; APPLICANT: Scheiflinger, Friedrich									
; APPLICANT: Kerschbaumer, Rndolf									
; APPLICANT: Falkner, Falko-Guenter									
; APPLICANT: Dornier, Friedrich									
; APPLICANT: Baxter Aktiengesellschaft									
; TITLE OF INVENTION: Factor IX/Factor IXa Activating Antibodies and Antibody									
; TITLE OF INVENTION: Derivatives									
; FILE REFERENCE: 20695C-005900US									
; CURRENT APPLICATION NUMBER: US/09/661,992B									
; CURRENT FILING DATE: 2000-09-14									
; PRIOR APPLICATION NUMBER: AT A157600									
; PRIOR FILING DATE: 1999-09-14									
; NUMBER OF SEQ ID NOS: 112									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 84									
; LENGTH: 249									
; TYPE: PRT									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: scFv from									
; OTHER INFORMATION: hybridoma cell line 193/K2									
US-09-661-992B-84									
Query Match 61.2%; Score 856; DB 3; Length 249;									
Best Local Similarity 65.8%; Pred. No. 9.9e-63;									
Matches 163; Conservative 38; Mismatches 43; Indels 6; Gaps 3;									
Qy	3	QVQLVDSGAEVKKPGASVKVSKASGYTFSNYGITWVROAPGQGLEWMGNN--THY	60						
Db	1	EVQLVESGGGLVQPGGSLKLSCAASGFTFSTYMTSWVRQTPKELRWATISSGSGSTYY	60						
Qy	61	AQKPGQGRVTMTDTSTSTAYMELRSLRSDDTAVVYICAR	---PRGDDPFWGQGITLVTVS	117					
Db	61	PDSVGRGFTISRDNKNTLYLQWNSLKSDDTAMYYCTRDGCHGYSFQYWGQGITLVTS	120						
Qy	118	SGGGSGGGSGGGSGGSEIVLTQSPSLSPVTLTGQPAISCRSSQNLVYSDGNLYLSPFOOR	177						
Db	121	SGGGSGGGSGRASGGGSGQIVLTQSPSLPVLGDPQASISCRSSQSIHNSGNITLYEYLQK	180						
Qy	178	PGQSPRLIYKVSNRDGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYVCMQGTWHPPTF	237						
Db	181	PGQSPKLLIYKVSNRSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYVCFQSGH-VPWTF	239						
Qy	238	GQGTKEVIEKR	247						

Db 240 GGGTKLEIKR 249

RESULT 14

US-09-726-219A-190

Sequence 190. Application US/09726219A

Patent No. 6806079

GENERAL INFORMATION:

APPLICANT: Cambridge Antibody Technology Limited

APPLICANT: Medical Research Council

APPLICANT: McCafferty, John

APPLICANT: Pope, Anthony

APPLICANT: Johnson, Kevin

APPLICANT: Hoogenboom, Hendricus

APPLICANT: Griffiths, Andrew

APPLICANT: Jackson, Ronald

APPLICANT: Holliger, Kasper

APPLICANT: Marks, James

APPLICANT: Clarkson, Timothy

APPLICANT: Chiswell, David

APPLICANT: Winter, Gregory

APPLICANT: Bonert, Timothy

TITLE OF INVENTION: Methods for Producing Members of Specific Binding Pairs

FILE REFERENCE: 213839-00013

CURRENT APPLICATION NUMBER: US/09/726,219A

CURRENT FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: GB 9015198.6

PRIOR FILING DATE: 1990-07-10

PRIOR APPLICATION NUMBER: GB 9022845.3

PRIOR FILING DATE: 1990-10-19

PRIOR APPLICATION NUMBER: GB 9022845.3

PRIOR FILING DATE: 1990-10-19

PRIOR APPLICATION NUMBER: GB 9024503.6

PRIOR FILING DATE: 1990-11-12

PRIOR APPLICATION NUMBER: GB 9104744.9

PRIOR FILING DATE: 1991-03-06

PRIOR APPLICATION NUMBER: GB 9110549.4

PRIOR FILING DATE: 1991-05-15

PRIOR APPLICATION NUMBER: PCT/GB91/01134

PRIOR FILING DATE: 1991-07-10

PRIOR APPLICATION NUMBER: US 07/971,857

PRIOR FILING DATE: 1993-01-08

PRIOR APPLICATION NUMBER: US 08/484,893

PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 272

SOFTWARE: PatentIn version 3.1

SEQ ID NO 190

LENGTH: 249

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: scFv form of the anti-oxazalone antibody NQ11

US-09-726-219A-190

Query Match 60.8%; Score 850; DB 2; Length 249;

Best Local Similarity 64.8%; Pred.No. 3.1e-62;

Matches 164; Conservative. 35; Mismatches 42; Indels 12; Gaps 4

Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGVTFSNYGITWTRQAPGQGLEWGWING---NT 58

Db 1 QVQLQESGGGLVQPGGSLRLSCATSGFTFSNYNGWVRQPEPKALEWLGSRVKNVGYTT 60

Qy 59 HYAQKFGQRTVMTTDTSTSTAYMELSLRSDTAVYICARDPRGDDE---PYWGQGLV 114

Db 61 EYSASVGRFTISRDNFQSIYLQINTLRDTSATYYCA---RGYDCAWFAFYWGQGLV 117

Qy 115 TVSSGGGGSGGGGGSGGSEIVLTQSPFLSPVTLTGQPASISCRSSQNLYVSDGNTYLSWF 174

Db 118 TVSSGGGGSGGGGGSGGSDIELTQTFLSLPVSIGDQASISCRSSQSIVHSNGNTYLEWY 177

Qy 175 QORPGQSPRLIYKVNRRDGPDRPFGSGSGGTDFTLKISRVEAEDIGVYVCMQGTWHP 234

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:26:47 ; Search time 179 Seconds

(without alignments)
675,414 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1398	100.0	261	6	US-11-056-825-2
2	1382	98.9	259	6	US-11-056-825-7
3	1135	81.2	253	3	US-09-880-748-1964
4	1135	81.2	253	4	US-10-293-418-1964
5	1135	81.2	253	6	US-11-054-515-1964
6	1135	81.2	253	6	US-11-266-444-1964
7	963	68.9	249	6	US-11-056-825-4
8	952	68.1	247	6	US-11-056-825-8
9	933.5	66.8	274	4	US-10-257-864A-92
10	933.5	66.8	274	5	US-10-645-085A-92
11	929.5	66.5	274	4	US-10-138-505-26
12	929.5	66.5	274	4	US-10-257-864A-90
13	929.5	66.5	274	4	US-10-399-518-119
14	929.5	66.5	274	5	US-10-399-585-118
15	929.5	66.5	274	5	US-10-645-085A-90
16	928.5	66.4	274	4	US-10-138-505-32
17	928.5	66.4	274	4	US-10-221-131-97
18	928.5	66.4	274	4	US-10-399-518-121
19	928.5	66.4	274	5	US-10-399-585-120
20	924.5	66.1	244	3	US-09-880-748-1991
21	924.5	66.1	244	4	US-10-293-418-1991
22	924.5	66.1	244	6	US-11-054-515-1991
23	924.5	66.1	244	6	US-11-266-444-1991
24	921.5	65.9	271	4	US-10-138-505-30
25	921.5	65.9	271	4	US-10-257-864A-91
26	921.5	65.9	271	4	US-10-221-131-96
27	921.5	65.9	271	4	US-10-399-518-120

28	921.5	65.9	271	5	US-10-399-585-119	Sequence 119, App
29	921.5	65.9	271	5	US-10-645-085A-91	Sequence 91, Appl
30	920.5	65.8	271	4	US-10-138-505-34	Sequence 34, Appl
31	920.5	65.8	271	4	US-10-257-864A-93	Sequence 93, Appl
32	920.5	65.8	271	4	US-10-221-131-98	Sequence 98, Appl
33	920.5	65.8	271	4	US-10-399-518-122	Sequence 122, App
34	920.5	65.8	271	5	US-10-399-585-121	Sequence 121, App
35	920.5	65.8	271	5	US-10-645-085A-93	Sequence 93, Appl
36	920.5	65.8	271	4	US-10-257-864A-96	Sequence 96, Appl
37	920.5	65.8	271	4	US-10-221-131-101	Sequence 101, App
38	920.5	65.8	271	4	US-10-399-518-125	Sequence 125, App
39	920.5	65.8	271	5	US-10-399-585-124	Sequence 124, App
40	920.5	65.8	271	5	US-10-645-085A-96	Sequence 96, Appl
41	919.5	65.8	271	4	US-10-239-656-47	Sequence 47, Appl
42	919.5	65.8	271	4	US-10-239-656-48	Sequence 48, Appl
43	919.5	65.8	271	4	US-10-239-656-49	Sequence 49, Appl
44	919.5	65.8	271	6	US-11-036-098-18	Sequence 18, Appl
45	914.5	65.4	271	3	US-09-968-851-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1

US-11-056-825-2

; Sequence 2, Application US/11056825

; Publication No. US20050255109A1

; GENERAL INFORMATION:

; APPLICANT: Janda, Kim D.

; APPLICANT: Saven, Alan

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS

; FILE REFERENCE: SCRP-0042

; CURRENT APPLICATION NUMBER: US/11/056,825

; CURRENT FILING DATE: 2005-02-11

; PRIOR APPLICATION NUMBER: US 60/626,726

; PRIOR FILING DATE: 2004-11-10

; PRIOR APPLICATION NUMBER: US 60/544,807

; PRIOR FILING DATE: 2004-02-13

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 2

; LENGTH: 261

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Synthetic Construct

US-11-056-825-2

Query Match 100.0%; Score 1398; DB 6; Length 261;
Best Local Similarity 100.0%; Pred. No. 2,7e-93;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAQVQLVQSGAEVKPGASVKSCASGTFNIGTITVTRAPQGLGWGNNGNTHY	60
Db	1	MAQVQLVQSGAEVKPGASVKSCASGTFNIGTITVTRAPQGLGWGNNGNTHY	60
QY	61	AQKFGQRTVMTDTSTAYMELRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGG	120
Db	61	AQKFGQRTVMTDTSTAYMELRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGG	120
QY	121	GGSGGGGGGGGSEIVLTQSLPVLTLQGPASISCRSSQNLVYSDGNTYLSWFQRPQG	180
Db	121	GGSGGGGGGGGSEIVLTQSLPVLTLQGPASISCRSSQNLVYSDGNTYLSWFQRPQG	180
QY	181	SPRLIYKVNDRSGVDPFRFGSGSGTDTLTKISRVEADIGVYICMQGTHWPPRTFGQ	240
Db	181	SPRLIYKVNDRSGVDPFRFGSGSGTDTLTKISRVEADIGVYICMQGTHWPPRTFGQ	240
QY	241	TKVEIKRGLGLVDYKDDDK	261
Db	241	TKVEIKRGLGLVDYKDDDK	261

```

RESULT 2
US-11-056-825-7
; Sequence 7, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SCRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: US 60/544,807
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-7

Query Match      98.9%; Score 1382; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 3.8e-92;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRQAPGGLEWMGWINNGNTHY 60
Db 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRQAPGGLEWMGWINNGNTHY 60

Qy 61 AOKQGRVTMTTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDDPYWGQGLTVTVSSGG 120
Db 61 AOKQGRVTMTTDTSTSTAYMELSLRSDDTAVVYCARDPRGDDDPYWGQGLTVTVSSGG 120

Qy 121 GSGGGGGGGGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVYSDGNTYLSWFOORPQG 180
Db 121 GSGGGGGGGGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVYSDGNTYLSWFOORPQG 180

Qy 181 SPRLLYKVSNRDGVDPFRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTTHWPPRTFGQG 240
Db 181 SPRLLYKVSNRDGVDPFRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTTHWPPRTFGQG 240

Qy 241 TKVEIKRGLGLVDYKDD 259
Db 241 TKVEIKRGLGLVDYKDD 259

RESULT 3
US-09-880-748-1964
; Sequence 1964, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/212,210
; NUMBER OF SEQ ID NOS: 3239

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1964

Query Match      81.2%; Score 1135; DB 3; Length 253;
Best Local Similarity 86.2%; Pred. No. 2.8e-74;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRQAPGGLEWMGWIN--NGNTHY 60
Db 1 QVQLQSGAEVKEPQASVKVSCKASGYTFTSYGISWRQAPGGLEWMGWINAYNGNTNY 60

Qy 61 AOKFQGRVTMTTDTSTSTAYMELSLRSDDTAVVYCARDPRGDD-----PYWGQGLTVT 115
Db 61 AOKLQGRVTMTTDTSTSTAYMELSLRSDDTAVVYCARPNPYVDSSEGGPFDYWGQGLTMT 120

Qy 116 VSSGGGGGGGGGGSG--EIVLTQSPSLPVLTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGSGSALDVVMTQSPSLPVLTLGQPASISCRSSQSLVYSDGNTYLNW 180

Qy 174 FOORPQSPRRLLYKVSNRDGVDPFRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTTHW 233
Db 181 FOORPQSPRRLLYKVSNRDGVDPFRFSGSGSGTDFTLKISRVEAEDVGVYCMQATRW- 239

Qy 234 PRTFGQGTKEIKR 247
Db 240 PRTFGQGTKEIKR 253

```

```

RESULT 4
US-10-293-418-1964
; Sequence 1964, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1964

Query Match      81.2%; Score 1135; DB 4; Length 253;
Best Local Similarity 86.2%; Pred. No. 2.8e-74;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRQAPGGLEWMGWIN--NGNTHY 60
Db 1 QVQLQSGAEVKEPQASVKVSCKASGYTFTSYGISWRQAPGGLEWMGWINAYNGNTNY 60

```

[illegible]

RESULT 5

```

US-11-054-515-1964
; Sequence 1964, Application US/11054515
; Publication No. US2005025532A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immune
; FILE REFERENCE: PF523P3
; CURRENT APPLICATION NUMBER: US/11/054,515
; CURRENT FILING DATE: 2005-02-10
; PRIOR APPLICATION NUMBER: 60/543,296
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/580,347
; PRIOR FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 10/293,418
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; Remaining prior Application data removed
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-054-515-1964

```

Query Match		81.2%;	Score 1135;	DB 6;	Length 253;	
Best Local Similarity		86.2%;	Pred. No. 2.8e-74;			
Matches 219; Conservative		15;	Mismatches 10;	Indels 10;	Gaps 4;	
<hr/>						
OY	3	QVQLVSGAEVKPGASVKVSCKASGYTFNYSIGITWRQAPGGGLEWMGWIN--NGNTHY	60			
Ddb	1	QVQLQQSGAEVKPEPGASVKVSCKASGYTFTSYGISWVRQAPGGGLEWMGWSA YNGNTNY	60			
<hr/>						
OY	61	AQKQFGRVTTMTDTSSTAYMELRSRSSDDTA VYYICARDPRGDDE-----PYKGQGTLVT	115			
Ddb	61	AQKLQGRVTMTDTSSTAYMELRSRSSDDTA VYYICARNFYDSDSEGFDFYWCQGHVT	120			
<hr/>						
OY	116	VSSGGGGSGGGSGGGGS--EIVLTQSPLPVT LGGQPASISCRSSONLYVSDGNTYLSW	173			
Ddb	121	VSSGGGGSGGGSGGGGSALDVVMTQSPLPVT LGGQPASISCRSSQSLVSDGNTYLNW	180			
<hr/>						
OY	174	FQRPGQSPPRLTYIKVSNRDSGVDPFSGSGSG TDFTLKISRVEABDIGVYVCMQGTHWP	233			

```

Db      181  FQBPQGSRRLLIYKVNNRDSGVDPFRFSGSGTGDTFLKISRVEADVGVIYCMQATRW- 239
QY      234  PRTEGQGTKEIKR 247
          | |||||: ||||
Db      240  PFTFGQGTKEIKR 253

```

RESULT 6
 US-11-266-444-1964
 ; Sequence 1964, Application US/11266444
 ; Publication No. US20060062789A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocytes
 ; FILE REFERENCE: PF523P1d1
 ; CURRENT APPLICATION NUMBER: US/11/266.444

CURRBL
; PRTOR

```

; PRIOR APPLICATION NUMBER: 037860, 748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-266-444-1964

Query Match      81.2%; Score 1135; DB 6; Length 253;
Best Local Similarity 86.2%; Pred. No. 2.8e-74;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4

QY      3  QVQLVSGAEVKKPGASVKVSCKASGYTFSSNYGILTWVRQAPQGGLRWGWIN--NGNTHY 60
DB      1  QVQLQSGAEVKKPGASVKVSCKASGYTFSSNYGILTWVRQAPQGGLRWGWIN--NGNTHY 60

QY      61  AQKPFQGRVTMTDTSTSTAYMELRSLRSDDTAVYYCARNPYYDSSEGFFDWGQGTWMT 115
DB      61  AQKLGQRVTMTDTSTSTAYMELRSLRSDDTAVYYCARNPYYDSSEGFFDWGQGTWMT 120

QY      116  VSSGGGGSGGGGGGGGG--EIVLTQSPFLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
DB      121  VSSGGGGSGGGGGGGGGGALDVVMTPSLPVLTLGQPASISCRSSQSLVYSDGNTYLNW 180

QY      174  FQRPQSPRLRIYKYSNRDSGVDPDRFSGSGSGTDFTLKISRVAEDIGVYYCMQGTWHP 233
DB      181  FQRPQSPRLRIYKYSNRDSGVDPDRFSGSGSGTDFTLKISRVAEDVGVIYCMQATRW- 239

QY      234  PRTFGQGTKEIKR 247
DB      240  PRTFGQGTKEIKR 253

RESULT 7
US-11-056-825-4
; Sequence 4, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825

```

```
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 4
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-4

Query Match      68.9%; Score 963; DB 6; Length 249;
Best Local Similarity 72.0%; Pred. No. 7,7e-62;
Matches 190; Conservative 29; Mismatches 27; Indels 18; Gaps 5;

Qy 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSNYGITWVRQAPGGQLEWMGWIN--NGNT 58
Db 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSNYGITWVRQAPGGQLEWMGWINPDSGDT 60

Qy 59 HYAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 118
Db 61 NSAQQFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 120

Qy 119 GGGSGGGSGGGGSEIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSWFOQRP 178
Db 61 NSAQQFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 120

Qy 121 G-----GGGSGSEIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQKP 167
Db 121 G-----GGGSGSEIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQKP 167

Qy 179 GQSPRLIYKVNRSRGVDPFRFSGSGGTDFTLKISRVEAEDIGVYYCMQ-GTHWPPRTF 237
Db 168 GQAPRLIYGASSRATGIPDRFSGSGGTDFTLKISRLEPEDFAVYYCQYGS--SPRTF 225

Qy 238 GQGTKVEIKRGLGLVDYKDDDK 261
Db 226 GQGTKVDIKRGLGLVDYKDDDK 249

RESULT 8
US-11-056-825-8
; Sequence 8, Application US/11056825
; Publication No. US20050255109A1
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SCRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 8
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-8

Query Match      68.1%; Score 952; DB 6; Length 247;
Best Local Similarity 71.8%; Pred. No. 4,8e-61;
Matches 188; Conservative 29; Mismatches 27; Indels 18; Gaps 5;

Qy 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSNYGITWVRQAPGGQLEWMGWIN--NGNT 58
Db 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSNYGITWVRQAPGGQLEWMGWINPDSGDT 60
```

```
Qy 59 HYAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 118
Db 61 NSAQQFQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSS 120

Qy 119 GGGSGGGSGGGGSEIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSWFOQRP 178
Db 121 G-----GGGSGSEIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAWYQKP 167

Qy 179 GQSPRLIYKVNRSRGVDPFRFSGSGGTDFTLKISRVEAEDIGVYYCMQ-GTHWPPRTF 237
Db 168 GQAPRLIYGASSRATGIPDRFSGSGGTDFTLKISRLEPEDFAVYYCQYGS--SPRTF 225

Qy 238 GQGTKVEIKRGLGLVDYKDDDK 259
Db 226 GQGTKVDIKRGLGLVDYKDDDK 247

RESULT 9
US-10-257-864A-92
; Sequence 92, Application US/10257864A
; Publication No. US20040058393A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, NAOSHI
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OH-EDA, MASAYOSHI
; APPLICANT: UNO, SHINSUKE
; APPLICANT: KIKUCHI, YASUFUMI
; APPLICANT: GOTEMBA-SHI, OHTOMO
; TITLE OF INVENTION: AGONIST ANTIBODIES
; FILE REFERENCE: 065678-0107
; CURRENT APPLICATION NUMBER: US/10/257,864A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: PCT/JP01/03288
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: JP2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP2000-321821
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: JP2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 92
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO: 24
US-10-257-864A-92

Query Match      66.8%; Score 933.5; DB 4; Length 274;
Best Local Similarity 69.8%; Pred. No. 1.2e-59;
Matches 185; Conservative 31; Mismatches 34; Indels 15; Gaps 5;

Qy 1 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSNYGITWVRQAPGGQLEWMGWI--NNGNT 58
Db 21 MAQVQLVQSGAEVKPKGASVKVSKASGYTFSNYGITWVRQAPGGQLEWMGWI--NNGNT 80

Qy 59 HYAQKFGQGRVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEY--WGQGLTVTV 116
Db 81 KYNEKPKRATLTSDKSTIATYMDLSSLASDGAIVYCA---RGYYTYDDWGQGLTVTV 137

Qy 117 SSGGGSGGGSGGGGSEIVLTQSPSLPVTLGQPASISCRSSQNLVYSDGNTYLSWFOQ 176
Db 138 SSGGGSGGGSGGGGSDVVMVTQSPSLPVLGQASISCRSSQSLVHSNGKTYLHWYVQ 197

Qy 177 RQGPSRRLIYKVNRSRGVDPFRFSGSGGTDFTLKISRVEAEDIGVYYCMQGTHWPPRT 236
Db 198 KPGQSPKLLIYKVNRSRGVDPFRFSGSGGTDFTLKISRVEAEDLGVYFCSQSTH-VPT 256
```

```
QY 237 FGGTKVEIKRGLGLVDYKDDDK 261
DB 257 FGGTKLEIK-----DYKDDDK 274

RESULT 10
US-10-645-085A-92
; Sequence 92, Application US/10645085A
; Publication No. US20040258684A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, NAOSHI
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OH-EDA, MASAYOSHI
; APPLICANT: UNO, SHINSUKE
; APPLICANT: KIKUCHI, YASUFUMI
; APPLICANT: OHTOMO, TOSHIHIKO
; TITLE OF INVENTION: AGONIST ANTIBODIES
; FILE REFERENCE: 065678-0108
; CURRENT APPLICATION NUMBER: US/10/645,085A
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/JP01/03288
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: JP 2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP 2000-321821
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: JP 2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 92
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO: 24
US-10-645-085A-92

Query Match 66.8%; Score 933.5; DB 5; Length 274;
Best Local Similarity 69.8%; Pred. No. 1.2e-59;
Matches 185; Conservative 31; Mismatches 34; Indels 15; Gaps 5;

QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITWVRQAPGQGLEWMGWI--NNGNT 58
DB 21 MAQVQLQSGAEIVKPGASVKMSCKASGYTFNHHVHWKQKPGQGLEWIGYIYPNDGT 80
QY 59 HYAQKFGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDEPY--WGQGTLLTV 116
DB 81 KYNEKFKDKATLTSDKSTTAYMDLSLASEDSAVYYCA---RGYYTYDDWGQGTLLTV 137
QY 117 SSGGGSGGGSGGGGGSEIVLTQSLPLVTLGQPASISCRSSONLVYSDGNTYLSWFPQ 176
DB 138 SSGGGSGGGSGGGGGSDVVMVTQSLPLVSLGDAQSISCRSSQSLVHNSKTYLHWYIQ 197
QY 177 RPGQSPRLIYKVNRSQVDPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTTHWPRT 236
DB 198 KPGQSPKLLIYKVNRFSGVPDRFSGSGSVTDFTLMISRVEAEDLGVYFCQSQSTH-VPYT 256

RESULT 11
US-10-138-505-26
; Sequence 26, Application US/10138505
; Publication No. US20030108546A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, Naoshi
; APPLICANT: UNO, Shinsuke
; APPLICANT: OH-EDA, Masayoshi
```

```
; APPLICANT: KIKUCHI, Yasufumi
; TITLE OF INVENTION: APOPTOSIS-INDUCING SINGLE-CHAIN FV
; FILE REFERENCE: 065678/0102
; CURRENT APPLICATION NUMBER: US/10/138,505
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US/09/523,095
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: JP 11-63557
; PRIOR FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-138-505-26

Query Match 66.5%; Score 929.5; DB 4; Length 274;
Best Local Similarity 69.1%; Pred. No. 2.2e-59;
Matches 183; Conservative 32; Mismatches 35; Indels 15; Gaps 5;

QY 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGITWVRQAPGQGLEWMGWI--NNGNT 58
DB 21 MAQVQLQSGSPDLVKPGASVKMSCKASGYTFVNHVHWKQKPGQGLEWIGYIYPNDGT 80
QY 59 HYAQKFGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDEPY--WGQGTLLTV 116
DB 81 KYNEKFKGKATLTSEKSSAAVMELSLASEDSAVYYCA---RGYYYSYDDWGQGTLLTV 137
QY 117 SSGGGSGGGSGGGGGSEIVLTQSLPLVTLGQPASISCRSSONLVYSDGNTYLSWFPQ 176
DB 138 SSGGGSGGGSGGGGGSDVVMVTQSLPLVSLGDAQSISCRSSQSLHSGKNTYLYQLQ 197
QY 177 RPGQSPRLIYKVNRSQVDPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTTHWPRT 236
DB 198 KPGQSPKLLIYKVNRFSGVPDRFSGSGSVTDFTLKISRVEAEDLGVYFCQSQSTH-VPYT 256

RESULT 12
US-10-257-864A-90
; Sequence 90, Application US/10257864A
; Publication No. US20040058393A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, NAOSHI
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OH-EDA, MASAYOSHI
; APPLICANT: UNO, SHINSUKE
; APPLICANT: KIKUCHI, YASUFUMI
; APPLICANT: GOTEMBA-SHI, OHTOMO
; TITLE OF INVENTION: AGONIST ANTIBODIES
; FILE REFERENCE: 065678-0107
; CURRENT APPLICATION NUMBER: US/10/257,864A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: PCT/JP01/03288
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: JP2000-115246
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: JP2000-321821
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: JP2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus musculus
```



```

; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO: 20
US-10-257-864A-90

```

Query Match	66.5%;	Score	929.5;	DB	4;	Length	274;
Best Local Similarity	69.1%;	Pred.	No. 2.2e-59;				
Matches	183;	Conservative	32;	Mismatches	35;	Indels	15;
Gaps	5;						

QY	1	MAQVOLVQSAEAVKPKGASVKVCKASGYTFSNYGITWVRQAPCGQLFWMGWI--NNGNT	58
DB	21	MAQVOLQOQSPDLVKPGASVKMSCKASGYTFVNRVHMHWKPKGQGLEWIGYIYPYNDGT	80
QY	59	HYAAQKFGRYVTMTDTSTAYMELRSLRSDDTAVVYCARDPRGDDEPY--WGQGLTVTV	116
DB	81	KYNEKPKGKATLLTSEKSSAAYMELSLASEDAVYCA---RGYYSYDDMGQGTITV	137
QY	117	SSGGSGSGSGSGGGGSGSEIVLTQSPISLPTVLGPASISCRSSONLYVSDGNTVLSLWFOQ	176
DB	138	SSGGSGSGSGSGGGSGSDVVMVTPTSLPVLGDAQASISCRSSQSLHLSKGNTVLQWYLO	197
QY	177	RPQSPRRLLIYKSNRSDGVPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTHTWPPRT	236
DB	198	KPGQSPKLLIYKSNRFGVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCSQSTH-VPYT	256
QY	237	FGQGTKEIKRGLGLVDYKDDDDK	261
DB	257	SGGQTKLEIK-----DYKDDDDK	274

```

RESULT 13
US-10-399-518-119
; Sequence 119, Application US/10399518
; Publication No. US20040091475A1
; GENERAL INFORMATION:
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: YABUTA, NAACHIRO
; APPLICANT: TSUNODA, HIROYUKI
; APPLICANT: ORITA, TETSURO
; TITLE OF INVENTION: DEGRADED TPO ACONIST ANTIBODY
; FILE REFERENCE: 065678/0111
; CURRENT APPLICATION NUMBER: US/10/399,518
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: PCT/JP01/03288
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: JP 2001-277314
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: JP 2000-321821
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-399-518-119

```

	Query Match	66.5%	Score 929.5	DB 4	Length 274
	Best Local Similarity	69.1%	Pred. No. 2.2e-59		
	Matches 183	Conservative 32	Mismatches 35	Indels 15	Gaps 5
Qy	1	MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSYNGYITWVRQAPGQGLEWMGWI--NNGNT	58		
Db	21	MAQVQLVQSGDPLVKPGASVKMSCKASGYTFVNHVMHWVQKPGQGLEWIGYIYPYNDGT	80		
Qy	59	HYAQKFGQVRVTMTTDTSTSTAYMELRLSRSDTAVYVCARDPRGDEBPY--WGQQTLLTV	116		
Db	81	KYNEKFGKATLISEKSSAAAYMELSLAEDSAVYCA---RGYYSYDDMGQGTLLTV	137		
Qy	117	SSGGSGSGGSGGGGGSEIVLTQPSLPLVT/LGPASISCRSSONLYVSDGNTYLSHFQQ	176		
Db	138	SSGGSGSGGSGGGGSDVMTQPTSLPSVSLGPOASISCRSSOSLHSGKNTYLOWYLO	197		

Qy	177	PCQSPRRLIYKYVSNRDSGVDPDRSGSGSGTDFTLKISRVEADLGVYVCMQCTHWPBPT	236
Db	198	KPGSPKLLIYKYVSNRDSGVDPDRSGSGSGTDFTLKISRVEADLGVYVFCQSTH-VPYT	256
Qy	237	FGQGTKEIIRKGLGLVDYKDDDDK	261
Db	257	SGGQTKLEIK-----DYKDDDDK	274

```

RESULT 14
US-1399-585-118
; Sequence 118, Application US/10399585
; Publication No. US2004024847A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, NAOSHI
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: UNO, SHINSUKE
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: YABUTA, NAOHRO
; APPLICANT: TSUNODA, HIROYUKI
; TITLE OF INVENTION: DEGRADED AGONIST ANTIBODY
; FILE REFERENCE: 065678/0112
; CURRENT APPLICATION NUMBER: US/10/399,585
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: PCT/JP01/092260
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/JP01/03298
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/JP01/01912
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: JP 2001-277314
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: JP 2000-321821
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: JP 2000-321822
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 118
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Mus sp.
; SS-10-399-585-118

```

Query Match	66.5%	Score	929.5	DB 5	Length	274			
Best Local Similarity	69.1%	Pred. No.	2.2e-59						
Matches	183	Conservative	32	Mismatches	35	Indels	15	Gaps	5
Qy	1	MAQVQLVQSGAEVKKPGQASVKVSCKASGYTFSNYGITVTRVQAFCQGLIEWMGWI--NNGNT	58						
Db	21	MAQVQLQQSGDLVLPKPGASVKMSCKASGYTFVNHVHMKVQKPGQGLEWIGYIYPYNDGT	80						
Qy	59	HYAQKFGQRVTMTDTSTAYMELSLRSDDTAVYVCARDPRGDDEPY--WGQGLITVTV	116						
Db	81	KYAEKFKGKATLTSEKSSSAAYMELSLASEDAVYCA---RGYYSYDDWGQGTTLTV	137						
Qy	117	SSGGSGSGGGSGGGGSEIVLTQSPISLPTLGPASISCRSSONLYVSDGNTVLSWFOQ	176						
Db	138	SSGGSGSGGGSGGGGSDVVMQTPTSLPVLSDQASISCRSSQSLHSGKNTYLOWYLO	197						
Qy	177	RPQGSPLRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYVCMQGTHTWPPRT	236						
Db	198	KPGQSPKLLIYKSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLVYVFCQSQSTH-VPYT	256						
Qy	237	FGQGTKEIKRGLGLVYIKDDDDK	261						
Db	257	SGGGTKEIK-----DYKDDDDK	274						

RESULT 15
US-10-645-085A-90
; Sequence 90, Application US/10645085A
; Publication No. US20040258684A1

GENERAL INFORMATION:
APPLICANT: FUKUSHIMA, NAOSHI
APPLICANT: TSUCHIYA, MASAYUKI
APPLICANT: OH-EDA, MASAYOSHI
APPLICANT: UNO, SHINSUKE
APPLICANT: KIKUCHI, YASUFUMI
APPLICANT: OHTOMO, TOSHIHIKO
TITLE OF INVENTION: AGONIST ANTIBODIES
FILE REFERENCE: 065678-0108
CURRENT APPLICATION NUMBER: US/10/645,085A
CURRENT FILING DATE: 2002-10-07
PRIOR APPLICATION NUMBER: PCT/JP01/03288
PRIOR FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: PCT/JP01/01912
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: JP 2000-115246
PRIOR FILING DATE: 2000-04-17
PRIOR APPLICATION NUMBER: JP 2000-321821
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: JP 2000-321822
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 90
LENGTH: 274
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO: 20
US-10-645-085A-90

Query Match 66.5%; Score 929.5; DB 5; Length 274;
Best Local Similarity 69.1%; Pred. No. 2.2e-59;
Matches 183; Conservative 32; Mismatches 35; Indels 15; Gaps 5;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGIIWVRQAPGQGLEWMGWI--NNGNT 58
Db 21 MAQVQLQSGPDLVKPGASVVMCKASGYTFVNHVMHWVKQKPGQGLEWIGYIYPNDGT 80
Qy 59 HYAOKFQGRVTMTTDTSTAYMELRSLSDDTAVYYCARDPRGDDEPY--WGQGTLLTV 116
Db 81 KYNEKFKGKATLTSEKSSAAIMELSSLSASDSAVYICA---RGYYSYDDWGQGTLLTV 137
Qy 117 SSGGGSGGGSGGGSGGSIIVLTQSPSLPVLTPGQPASISCRSSQNLVYSDGNTYLSWFQQ 176
Db 138 SSGGGSGGGSGGGSGGSDVMTQTPLSLPVSLGDAQASISCRSSQSLHSGKNTYLGWYLQ 197
Qy 177 RPOGSPRLIYKVNRSRGVDPDFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPRT 236
Db 198 KPGQSPKLLIYKVNRFSGVDPDFSGSGGTDTFTLKISRVEAEDIGVYFCQSQTH-VPYT 256
Qy 237 FQGTKEIKRGLGLVDYKDDDK 261
Db 257 SGGGTLKLEIK-----DYKDDDK 274

Search completed: August 9, 2006, 13:30:14
Job time : 180 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:27:22 ; Search time 33 Seconds

(without alignments)
532.377 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 239914 seqs, 67312017 residues

Total number of hits satisfying chosen parameters: 239914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	943	67.5	256	7	US-11-337-300-278
2	863	61.7	252	7	US-11-337-300-290
3	846	60.5	697	7	US-11-155-444-2
4	846	60.5	697	7	US-11-156-109-2
5	846	60.5	731	7	US-11-155-444-6
6	846	60.5	731	7	US-11-156-109-6
7	813	58.2	252	6	US-10-684-237-4
8	810	57.9	247	7	US-11-337-300-65
9	799.5	57.2	244	6	US-10-539-402-32
10	791	56.6	254	7	US-11-337-300-216
11	786	56.2	254	7	US-11-337-300-284
12	783.5	56.0	255	7	US-11-337-300-242
13	783	56.0	254	7	US-11-337-300-218
14	782	55.9	254	7	US-11-337-300-276
15	780	55.8	254	7	US-11-337-300-274
16	779	55.7	254	7	US-11-337-300-240
17	777	55.6	247	7	US-11-337-300-36
18	774	55.4	254	7	US-11-337-300-214
19	774	55.4	269	6	US-10-539-402-1
20	772.5	55.3	253	7	US-11-337-300-244
21	768	54.9	252	7	US-11-337-300-264
22	768	54.9	252	7	US-11-337-300-288
23	767	54.9	256	7	US-11-337-300-286
24	765	54.7	252	7	US-11-337-300-232
25	765	54.7	252	7	US-11-337-300-270

26	765	54.7	252	7	US-11-337-300-272	Sequence 272, Appl
27	764.5	54.7	242	7	US-11-183-325-20	Sequence 20, Appl
28	764.5	54.7	255	7	US-11-337-300-238	Sequence 238, Appl
29	764	54.6	252	7	US-11-337-300-252	Sequence 252, Appl
30	763.5	54.6	279	6	US-10-968-757-2	Sequence 2, Appl
31	763	54.6	252	7	US-11-337-300-266	Sequence 266, Appl
32	762.5	54.5	242	7	US-11-183-325-18	Sequence 18, Appl
33	761	54.4	252	7	US-11-337-300-228	Sequence 228, Appl
34	761	54.4	252	7	US-11-337-300-280	Sequence 280, Appl
35	759.5	54.3	248	6	US-10-539-402-17	Sequence 17, Appl
36	759	54.3	252	7	US-11-337-300-226	Sequence 226, Appl
37	758.5	54.3	254	7	US-11-075-891-6	Sequence 6, Appl
38	758	54.2	252	7	US-11-337-300-256	Sequence 256, Appl
39	757	54.1	252	7	US-11-337-300-254	Sequence 254, Appl
40	756	54.1	252	7	US-11-337-300-230	Sequence 230, Appl
41	755	54.0	260	6	US-10-991-309B-160	Sequence 160, Appl
42	754.5	54.0	254	7	US-11-075-891-8	Sequence 8, Appl
43	754	53.9	252	7	US-11-337-300-268	Sequence 268, Appl
44	753.5	53.9	251	7	US-11-337-300-212	Sequence 212, Appl
45	753	53.9	249	7	US-11-337-300-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1

US-11-337-300-278

; Sequence 278, Application US/11337300

; Publication No. US20060121580A1

; GENERAL INFORMATION:

; APPLICANT: Crucell Holland B.V.

; APPLICANT: ter Meulen, Jan H.

; APPLICANT: De Kruif, Cornelis A.

; APPLICANT: van den Brink, Edward N.

; APPLICANT: Goudsmit, Jaap

; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof

; FILE REFERENCE: 0091 WO 00 ORD

; CURRENT APPLICATION NUMBER: US/11/337,300

; CURRENT FILING DATE: 2006-01-20

; NUMBER OF SEQ ID NOS: 478

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 278

; LENGTH: 256

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: SC03-053

US-11-337-300-278

Query Match

Best Local Similarity 67.5%; Score 943; DB 7; Length 256;

Matches 190; Conservative 19; Mismatches 35; Indels 12; Gaps 5;

QY	1	MAQVQLVQSGAEVKKPGASVKVSKASGYTFSNYGTVWVROAPGQGLEWVGWIN--NGNT 58
DB	2	MAQVQLVQSGAEVKKPGASVMSVSKASGYTFSNYAMHVVWVROGQRLWVGWINADRGQT 61
QY	59	HYAOKFQGRVTMTDTSTAYMELSLRSDDTAVVYCARDP-----RGDDEPYWGQT 112
DB	62	KYSQKFGKVTITGDTASTAYMDLSLRSEDFAVYICARGTGYLRSYHGMD--VMGQGT 119
QY	113	LVTVSSGGSGGGGGGGGGS-EIVLTQSPFLSLPVTLPQSPASISCRSSQNLVYSDGNTYL 171
DB	120	TVTVSSGGSGGGSGGGSGTGGTDDVMTQSPPSLPVTLPGEFASISCRSSQILHSNGNYL 179
QY	172	SNFQRPQGSPPRLIYKVSNRDGVDRFSGSGGTDFTLKISRVEAEDTGVYYCMQGT 231
DB	180	DWYLOKPGSPQLIYLGSNRASGVDRFSGSGGTDFTLKISRVEAEDGVYYCMQALQ 239
QY	232	WPPRTFGQGTKEIKR 247
DB	240	-TPLTFGQGTLEIKR 254

RESULT 2
US-11-337-300-290
; Sequence 290, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudamit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT APPLICATION NUMBER: US/11/337,300
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 290
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: SC03-059
US-11-337-300-290

Query Match 61.7%; Score 863; DB 7; Length 252;
Best Local Similarity 68.6%; Pred. No. 1.6e-56;
Matches 175; Conservative 26; Mismatches 40; Indels 14; Gaps 6;

QY 1 MAQVQLVQSGAEVKKPKQASVKVSKAS-GYTFESNYGITVVRQAPGGQLEWGMWIN--NGN 57
DB 2 MAQVQLVQSGAEVKKPGSSVKVSKASGGGVFRFYAINVVRQAPGGQLEWGMINPSGGS 61
QY 58 THYAKRFGQGRVTMTDITSTAYMELRSLRSDDTAVVYCARDPRGDE----PYWQGTL 113
DB 62 TSAQKRFQGRVLTITDITSTVYMWELSSLSSEDTAVVYCARPGGTSRGYMDVWGKGT 121
QY 114 VTSSGGSGGGGGGGGGGGG-EIVLTQSPSLPVTILGQSPASISCRSSQNLVSDGNTYLS 172
DB 122 VTSSGGSGGGGGGGGGGGGTEIVLTQSPATLSLSPGERATLSCRASQSV----SSYLA 176
QY 173 WFOQRGQSPRLIYKVNRRDSGVPRFSGSGSGTDFTLTKSRVEAEIDGVVYCMQGTW 232
DB 177 WYQKFGQAPRLIYDASNRATGIPARFSGSGSGTDFTLTISLSEPEDFAVYVYCCQQRSNW 236
QY 233 PRRTFGGTKEIKR 247
DB 237 PP-AFGGTKEIKR 250

RESULT 3
US-11-155-444-2
; Sequence 2, Application US/11155444
; Publication No. US20060104971A1
; GENERAL INFORMATION:
; APPLICANT: GARBNER, ELLEN
; APPLICANT: BAILLY, VERONIQUE
; APPLICANT: BROWNING, JEFFREY L.
; TITLE OF INVENTION: MULTIVALENT LYMPHOTOXIN BETA RECEPTOR AGONISTS AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; FILE REFERENCE: BGN168CN
; CURRENT APPLICATION NUMBER: US/11/155,444
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/US03/041393
; PRIOR FILING DATE: 2003-12-22
; PRIOR APPLICATION NUMBER: 60/435,154
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/435,185
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 2
; LENGTH: 697


```
; OTHER INFORMATION: SC03-021
US-11-337-300-216

Query Match      56.6%; Score 791; DB 7; Length 254;
Best Local Similarity 63.2%; Pred. No. 3.1e-51;
Matches 160; Conservative 33; Mismatches 52; Indels 8; Gaps 6;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRAPGQGLEWMGWINGN--T 58
Db 2 MAQVQLVQSGTEVKKPGESLKISKSGSGYFTYWGVRQMPGKGLWMIYFGDSET 61
Qy 59 HYAQKFGQRTMTTDTSTAYMELRSDDTAVYVCARDPRGDDP--YWGQGLTVTV 116
Db 62 RYSPFQGVQVITISADKSNINTAYLQWSSLKASDTAIYCA--GGSGISTPMDVMWQGTITV 120
Qy 117 SSGGGGGGGGGGGGG--EIVLTQSPSLPVTLPQASISCRSSQNLVYSDGN--TYLSWF 174
Db 121 SSGTGGSGGTGGTGGTDDVMTQSPDSLAVSLGERATINCKSSQSVLHSSNNKNYLAWY 180
Qy 175 QORPGQSPRLIYKVNRSQVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPP 234
Db 181 QORPGQSPRLIYKVNRSQVPRFSGSGGTDTFTLTISLQAEADVAVYCCQYISTP 240
Qy 235 RTFGQGTKEIKR 247
Db 241 -TFGQGTKEIKR 252

RESULT 11
US-11-337-300-284
; Sequence 284, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 284
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURES:
; OTHER INFORMATION: SC03-056
US-11-337-300-284

Query Match      56.2%; Score 786; DB 7; Length 254;
Best Local Similarity 62.5%; Pred. No. 7.3e-51;
Matches 158; Conservative 36; Mismatches 51; Indels 8; Gaps 6;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRAPGQGLEWMGWINGN--T 58
Db 2 MAEVQLVQSGTEVKKPGESLKISKSGSGYFTYWGVRQMPGKGLWMIYFGDSET 61
Qy 59 HYAQKFGQRTMTTDTSTAYMELRSDDTAVYVCARDPRGDDP--YWGQGLTVTV 116
Db 62 RYSPFQGVQVITISADKSNINTAYLQWSSLKASDTAIYCA--GGSGISTPMDVMWQGTITV 120
Qy 117 SSGGGGGGGGGGGGG--EIVLTQSPSLPVTLPQASISCRSSQNLVYSDGN--TYLSWF 174
Db 121 SSGTGGSGGTGGTGGTDDVMTQSPDSLAVSLGERATINCKSSQSVLHSSNNKNYLAWY 180
Qy 175 QORPGQSPRLIYKVNRSQVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPP 234
Db 181 QORPGQSPRLIYKVNRSQVPRFSGSGGTDTFTLTISLQAEADVAVYCCQYISTP 239
Qy 235 RTFGQGTKEIKR 247

; OTHER INFORMATION: SC03-021
US-11-337-300-216

Query Match      56.0%; Score 783.5; DB 7; Length 255;
Best Local Similarity 62.7%; Pred. No. 1.1e-50;
Matches 158; Conservative 34; Mismatches 53; Indels 7; Gaps 5;

Qy 1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRAPGQGLEWMGWINGN--T 58
Db 2 MAEVQLVQSGAEVKKPGESLKISKSGSGYFTYWGVRQMPGKGLWMIYFGDSET 61
Qy 59 HYAQKFGQRTMTTDTSTAYMELRSDDTAVYVCARDPRGDDP--YWGQGLTVTV 116
Db 62 RYSPFQGVQVITISADKSNINTAYLQWSSLKASDTAIYCA--GGSGISTPMDVMWQGTITV 120
Qy 117 SSGGGGGGGGGGGGG--EIVLTQSPSLPVTLPQASISCRSSQNLVYSDGN--TYLSWF 174
Db 121 SSGTGGSGGTGGTGGTDDVMTQSPDSLAVSLGERATINCKSSQSVLHSSNNKNYLAWY 180
Qy 175 QORPGQSPRLIYKVNRSQVPRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTTHWPP 234
Db 181 QORPGQSPRLIYKVNRSQVPRFSGSGGTDTFTLTISLQAEADVAVYCCQYISTP 240
Qy 235 RTFGQGTKEIKR 246
Db 241 YSFGQGTKEIKR 252

RESULT 13
US-11-337-300-218
; Sequence 218, Application US/11337300
; Publication No. US20060121580A1
; GENERAL INFORMATION:
; APPLICANT: Crucell Holland B.V.
; APPLICANT: ter Meulen, Jan H.
; APPLICANT: De Kruif, Cornelis A.
; APPLICANT: van den Brink, Edward N.
; APPLICANT: Goudsmit, Jaap
; TITLE OF INVENTION: Binding molecules against SARS-coronavirus and uses thereof
; FILE REFERENCE: 0091 WO 00 ORD
; CURRENT FILING DATE: 2006-01-20
; NUMBER OF SEQ ID NOS: 478
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial sequence
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:15:37 ; Search time 603 Seconds
(without alignments)
660.921 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 8366291 seqs, 1526956180 residues

Total number of hits satisfying chosen parameters: 8366291

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main.*

1: /EMC_Celerra_SIDS3/ptodata/2/paa/US066 COMB.pap.*
2: /EMC_Celerra_SIDS3/ptodata/2/paa/US066 COMB.pap.*
3: /EMC_Celerra_SIDS3/ptodata/2/paa/US073 COMB.pap.*
4: /EMC_Celerra_SIDS3/ptodata/2/paa/US073 COMB.pap.*
5: /EMC_Celerra_SIDS3/ptodata/2/paa/US075 COMB.pap.*
6: /EMC_Celerra_SIDS3/ptodata/2/paa/US076 COMB.pap.*
7: /EMC_Celerra_SIDS3/ptodata/2/paa/US077 COMB.pap.*
8: /EMC_Celerra_SIDS3/ptodata/2/paa/US078 COMB.pap.*
9: /EMC_Celerra_SIDS3/ptodata/2/paa/US079 COMB.pap.*
10: /EMC_Celerra_SIDS3/ptodata/2/paa/US080 COMB.pap.*
11: /EMC_Celerra_SIDS3/ptodata/2/paa/US081 COMB.pap.*
12: /EMC_Celerra_SIDS3/ptodata/2/paa/US082 COMB.pap.*
13: /EMC_Celerra_SIDS3/ptodata/2/paa/US083 COMB.pap.*
14: /EMC_Celerra_SIDS3/ptodata/2/paa/US084 COMB.pap.*
15: /EMC_Celerra_SIDS3/ptodata/2/paa/US085 COMB.pap.*
16: /EMC_Celerra_SIDS3/ptodata/2/paa/US086 COMB.pap.*
17: /EMC_Celerra_SIDS3/ptodata/2/paa/US087 COMB.pap.*
18: /EMC_Celerra_SIDS3/ptodata/2/paa/US088 COMB.pap.*
19: /EMC_Celerra_SIDS3/ptodata/2/paa/US089 COMB.pap.*
20: /EMC_Celerra_SIDS3/ptodata/2/paa/US090 COMB.pap.*
21: /EMC_Celerra_SIDS3/ptodata/2/paa/US091 COMB.pap.*
22: /EMC_Celerra_SIDS3/ptodata/2/paa/US092 COMB.pap.*
23: /EMC_Celerra_SIDS3/ptodata/2/paa/US093 COMB.pap.*
24: /EMC_Celerra_SIDS3/ptodata/2/paa/US094 COMB.pap.*
25: /EMC_Celerra_SIDS3/ptodata/2/paa/US095 COMB.pap.*
26: /EMC_Celerra_SIDS3/ptodata/2/paa/US096 COMB.pap.*
27: /EMC_Celerra_SIDS3/ptodata/2/paa/US097 COMB.pap.*
28: /EMC_Celerra_SIDS3/ptodata/2/paa/US098 COMB.pap.*
29: /EMC_Celerra_SIDS3/ptodata/2/paa/US099 COMB.pap.*
30: /EMC_Celerra_SIDS3/ptodata/2/paa/US100 COMB.pap.*
31: /EMC_Celerra_SIDS3/ptodata/2/paa/US101 COMB.pap.*
32: /EMC_Celerra_SIDS3/ptodata/2/paa/US102 COMB.pap.*
33: /EMC_Celerra_SIDS3/ptodata/2/paa/US103 COMB.pap.*
34: /EMC_Celerra_SIDS3/ptodata/2/paa/US104 COMB.pap.*
35: /EMC_Celerra_SIDS3/ptodata/2/paa/US105 COMB.pap.*
36: /EMC_Celerra_SIDS3/ptodata/2/paa/US106 COMB.pap.*
37: /EMC_Celerra_SIDS3/ptodata/2/paa/US107 COMB.pap.*
38: /EMC_Celerra_SIDS3/ptodata/2/paa/US108 COMB.pap.*
39: /EMC_Celerra_SIDS3/ptodata/2/paa/US109 COMB.pap.*
40: /EMC_Celerra_SIDS3/ptodata/2/paa/US110 COMB.pap.*
41: /EMC_Celerra_SIDS3/ptodata/2/paa/US111 COMB.pap.*
42: /EMC_Celerra_SIDS3/ptodata/2/paa/US112 COMB.pap.*
43: /EMC_Celerra_SIDS3/ptodata/2/paa/US113 COMB.pap.*

44: /EMC_Celerra_SIDS3/ptodata/2/paa/US114 COMB.pap.*
45: /EMC_Celerra_SIDS3/ptodata/2/paa/US600 COMB.pap.*
46: /EMC_Celerra_SIDS3/ptodata/2/paa/US601 COMB.pap.*
47: /EMC_Celerra_SIDS3/ptodata/2/paa/US602 COMB.pap.*
48: /EMC_Celerra_SIDS3/ptodata/2/paa/US603 COMB.pap.*
49: /EMC_Celerra_SIDS3/ptodata/2/paa/US604 COMB.pap.*
50: /EMC_Celerra_SIDS3/ptodata/2/paa/US605 COMB.pap.*
51: /EMC_Celerra_SIDS3/ptodata/2/paa/US606 COMB.pap.*
52: /EMC_Celerra_SIDS3/ptodata/2/paa/US607 COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1398	100.0	261	1	PCT-US05-04612-2
2	1398	100.0	261	40	US-11-056-825-2
3	1382	98.9	259	1	PCT-US05-04612-7
4	1382	98.9	259	40	US-11-056-825-7
5	1135	81.2	253	1	PCT-US01-19110-1964
6	1135	81.2	253	1	PCT-US02-36496-1964
7	1135	81.2	253	28	US-09-880-748-1964
8	1135	81.2	253	32	US-10-293-418-1964
9	1135	81.2	253	40	US-11-054-515-1964
10	1135	81.2	253	42	US-11-266-444-1964
11	1135	81.2	253	52	US-60-725-626-1964
12	1135	81.2	253	52	US-60-735-988-1964
13	1135	81.2	253	52	US-60-776-665-1964
14	1011.5	72.4	242	22	US-09-202-000-14
15	988.5	70.7	252	37	US-10-762-593-589
16	988.5	70.7	252	38	US-10-854-439-589
17	983.5	70.4	286	35	US-10-535-764-184
18	982.5	70.3	251	1	PCT-US00-19843-11
19	982.5	70.3	251	30	US-10-031-783-11
20	978.5	70.0	247	25	US-09-538-038-2343
21	978.5	70.0	247	25	US-09-538-038A-2343
22	963	68.9	249	1	PCT-US05-04612-4
23	963	68.9	249	40	US-11-056-825-4
24	952.5	68.1	524	35	US-10-551-504-293
25	952	68.1	247	1	PCT-US05-04612-8
26	952	68.1	247	40	US-11-056-825-8
27	949.5	67.9	524	35	US-10-551-504-287
28	947.5	67.8	524	35	US-10-551-504-2
29	943	67.5	256	43	US-11-337-300-278
30	939	67.2	281	35	US-10-535-764-186
31	937.5	67.1	524	35	US-10-551-504-260
32	933.5	66.8	274	32	US-10-257-864A-92
33	933.5	66.8	274	36	US-10-645-085A-92
34	933.5	66.8	524	35	US-10-551-504-254
35	929.5	66.5	274	25	US-09-523-095A-26
36	929.5	66.5	274	31	US-10-138-505-26
37	929.5	66.5	274	32	US-10-257-864A-90
38	929.5	66.5	274	33	US-10-399-518-119
39	929.5	66.5	274	33	US-10-399-585-118
40	929.5	66.5	274	36	US-10-645-085A-90
41	928.5	66.4	274	25	US-09-523-095A-32
42	928.5	66.4	274	31	US-10-138-505-32
43	928.5	66.4	274	32	US-10-221-131-97
44	928.5	66.4	274	33	US-10-399-518-121
45	928.5	66.4	274	33	US-10-399-585-120

ALIGNMENTS

RESULT 1
PCT-US05-04612-2
; Sequence 2, Application PC/TUS0504612
; GENERAL INFORMATION:

APPLICANT: Felding-Habermann, Brunhilde
APPLICANT: Janda, Kim D.
APPLICANT: Saven, Alan
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
FILE REFERENCE: SCRP-0043
CURRENT APPLICATION NUMBER: PCT/US05/04612
CURRENT FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: US 11/056,825
PRIOR FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: US 60/626,726
PRIOR FILING DATE: 2004-11-10
PRIOR APPLICATION NUMBER: US 60/544,807
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent in version 3.3
SEQ ID NO 2
LENGTH: 261
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic Construct
PCT-US05-04612-2

Query Match 100.0%; Score 1398; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-107;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
QY 181 SPRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
DB 181 SPRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDK 261

RESULT 2
US-11-056-825-2
Sequence 2, Application US/11056825
GENERAL INFORMATION:
APPLICANT: Felding-Habermann, Brunhilde
APPLICANT: Janda, Kim D.
APPLICANT: Saven, Alan
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
FILE REFERENCE: SCRP-0042
CURRENT APPLICATION NUMBER: US/11/056,825
CURRENT FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: US 60/626,726
PRIOR FILING DATE: 2004-11-10
PRIOR APPLICATION NUMBER: US 60/544,807
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent in version 3.3
SEQ ID NO 2
LENGTH: 261
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic Construct
US-11-056-825-2

QY 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
QY 181 SPRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
DB 181 SPRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDK 261

Query Match 100.0%; Score 1398; DB 40; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-107;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
QY 181 SPRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
DB 181 SPRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDK 261
DB 241 TKVEIKRGLGLVDYKDDDK 261

RESULT 3
PCT-US05-04612-7
Sequence 7, Application PC/TUS0504612
GENERAL INFORMATION:
APPLICANT: Felding-Habermann, Brunhilde
APPLICANT: Janda, Kim D.
FILE REFERENCE: SCRP-0043
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
CURRENT APPLICATION NUMBER: PCT/US05/04612
CURRENT FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: US 11/056,825
PRIOR FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: US 60/626,726
PRIOR FILING DATE: 2004-11-10
PRIOR APPLICATION NUMBER: US 60/544,807
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent in version 3.3
SEQ ID NO 7
LENGTH: 259
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic Construct
PCT-US05-04612-7

Query Match 98.9%; Score 1382; DB 1; Length 259;
Best Local Similarity 99.6%; Pred. No. 5.6e-106;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
DB 1 MAQVQLVQSGAEVKPKGASVKSCASGYTFSPNYGITTWRQAPGQGLEWMGWINNNGNTHY 60
QY 61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
DB 61 AQKFGQRTVTMTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
QY 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
DB 121 GSGGGGGGGGGSEIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSWFQORPGQ 180
QY 181 SPRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
DB 181 SPRLIYKVSNRDSGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTWHPRTFGQ 240
QY 241 TKVEIKRGLGLVDYKDDDD 259

```
Db      241 TKVEIKRGLGLVDYKDD 259
|||||
RESULT 4
US-11-056-825-7
; Sequence 7, Application US/11056825
; GENERAL INFORMATION:
; APPLICANT: Felding-Habermann, Brunhilde
; APPLICANT: Janda, Kim D.
; APPLICANT: Saven, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF METASTASIS
; FILE REFERENCE: SGRP-0042
; CURRENT APPLICATION NUMBER: US/11/056,825
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/626,726
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/544,807
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.3
; SEQ ID NO. 7
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-056-825-7

Query Match      98.9%; Score 1382; DB 40; Length 259;
Best Local Similarity 99.6%; Pred. No. 5.6e-106;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MAQVQLVQSGAEVKPGASVKSCASGYTFSNYGITWVRQAPGQGLEWMGWINNNTY 60
Db      1 MAQVQLVQSGAEVKPGASVKSCASGYTFSNYGITWVRQAPGQGLEWMGWINNNTY 60
Qy      61 AQKFGQRTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Db      61 AQKFGQRTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDPEYWGQGLTVTVSSGG 120
Qy      121 GSGGGGGGGGGSEIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180
Db      121 GSGGGGGGGGGSEIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOQRPQG 180
Qy      181 SPRLIYKVNRRDGVPRDFSGSGSGTDTLTKISRVEADIGVYVCMQGTWHPRTFGQ 240
Db      181 SPRLIYKVNRRDGVPRDFSGSGSGTDTLTKISRVEREDIGVYVCMQGTWHPRTFGQ 240
Qy      241 TKVEIKRGLGLVDYKDD 259
Db      241 TKVEIKRGLGLVDYKDD 259

RESULT 5
PCT-US01-19110-1964
; Sequence 1964, Application PC/TUS0119110
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind Blys
; FILE REFERENCE: PF523PCT
; CURRENT APPLICATION NUMBER: PCT/US01/19110
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25

Db      241 TKVEIKRGLGLVDYKDD 259
|||||
US-11-056-825-2
```

```
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-19110-1964

Query Match      81.2%; Score 1135; DB 1; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy      3 QVQLVQSGAEVKPGASVKSCASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNTHY 60
Db      1 QVQLQQSGAEVKPGASVKSCASGYTFTSYGISWVRQAPGQGLEWMGWSAYNGNTNY 60
Qy      61 AQKFGQRTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGLTV 115
Db      61 AQKLGQRTMTTDTSTSTAYMELSLRSDDTAVYYCARPNYYDSSEGFDPYWGQGLTV 120
Qy      116 VSSGGGGGGGGGG--EIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 173
Db      121 VSSGGGGGGGGGG--EIVLTQSPSLPVTILGQPASISCRSSQSLVYSDGNTYLNW 180
Qy      174 FQORFGQSPRLIYKVNRRDGVPRDFSGSGSGTDTLTKISRVEADIGVYVCMQGTWHP 233
Db      181 FQORFGQSPRLIYKVNRRDGVPRDFSGSGSGTDTLTKISRVEADVGVYVCMQATRW- 239
Qy      234 PRTFGQGTKEIKR 247
Db      240 PRTFGQGTKEIKR 253

RESULT 6
PCT-US02-36496-1964
; Sequence 1964, Application PC/TUS0236496
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind Blys
; FILE REFERENCE: PF523PCT2
; CURRENT APPLICATION NUMBER: PCT/US02/36496
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-36496-1964

Query Match      81.2%; Score 1135; DB 1; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy      3 QVQLVQSGAEVKPGASVKSCASGYTFSNYGITWVRQAPGQGLEWMGWIN--NGNTHY 60
Db      1 QVQLQQSGAEVKPGASVKSCASGYTFTSYGISWVRQAPGQGLEWMGWSAYNGNTNY 60
Qy      61 AQKFGQRTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDD-----PYWGQGLTV 115
Db      61 AQKLGQRTMTTDTSTSTAYMELSLRSDDTAVYYCARPNYYDSSEGFDPYWGQGLTV 120
Qy      116 VSSGGGGGGGGGG--EIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 173
Db      121 VSSGGGGGGGGGG--EIVLTQSPSLPVTILGQPASISCRSSQSLVYSDGNTYLNW 180
Qy      174 FQORFGQSPRLIYKVNRRDGVPRDFSGSGSGTDTLTKISRVEADIGVYVCMQGTWHP 233
Db      181 FQORFGQSPRLIYKVNRRDGVPRDFSGSGSGTDTLTKISRVEADVGVYVCMQATRW- 239
```

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

RESULT 7

US-09-880-748-1964
; Sequence 1964, Application US/09880748
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1964

Query Match 81.2%; Score 1135; DB 28; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGGQGLEMMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPCGASVKVSCKASGYTFTSYGISWVRQAPGQGLEMMGWISAYNGNTNY 60
Qy 61 AOKFQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 115
Db 61 AOKLQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARNPYDSSGFFDYWGQGTMT 120
Qy 116 VSSGGSGGGGGSGGGGS--EIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGSGGGGGSGGGGSALDVMTQSPSLPVTLCQPASISCRSSQSLVYSDGNTYLNW 180
Qy 174 FQORPGQSPRLIYKVSNRDGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQORPGQSPRLIYKVSNRDGVDPDRFSGSGSGTDFTLKISRVEAEDVGYYCMQATRW- 239
Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

RESULT 8

US-10-293-418-1964
; Sequence 1964, Application US/10293418
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1964

Query Match 81.2%; Score 1135; DB 32; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
Qy 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGGQGLEMMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPCGASVKVSCKASGYTFTSYGISWVRQAPGQGLEMMGWISAYNGNTNY 60
Qy 61 AOKFQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 115
Db 61 AOKLQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARNPYDSSGFFDYWGQGTMT 120
Qy 116 VSSGGSGGGGGSGGGGS--EIVLTQSPSLPVTLCQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGSGGGGGSGGGGSALDVMTQSPSLPVTLCQPASISCRSSQSLVYSDGNTYLNW 180
Qy 174 FQORPGQSPRLIYKVSNRDGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQORPGQSPRLIYKVSNRDGVDPDRFSGSGSGTDFTLKISRVEAEDVGYYCMQATRW- 239
Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

RESULT 9

US-11-054-515-1964
; Sequence 1964, Application US/11054515
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P3
; CURRENT APPLICATION NUMBER: US/11/054,515
; CURRENT FILING DATE: 2005-02-10
; PRIOR APPLICATION NUMBER: 60/543,296
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/580,347
; PRIOR FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 10/293,418
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964

```
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-054-515-1964

Query Match      81.2%; Score 1135; DB 40; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGIVTWVRQAPQGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPKGASVKVSKASGYTFTSYGISWVRQAPQGQGLEWMGWSAYNGNTNY 60

Qy 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVMTTDTSTAYMELSLRSDDTAVYYCARNPYTDSSEGFDFYWGQGLT 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180

Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

Query Match      81.2%; Score 1135; DB 42; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGIVTWVRQAPQGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPKGASVKVSKASGYTFTSYGISWVRQAPQGQGLEWMGWSAYNGNTNY 60

Qy 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVMTTDTSTAYMELSLRSDDTAVYYCARNPYTDSSEGFDFYWGQGLT 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180

Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

Query Match      81.2%; Score 1135; DB 44; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGIVTWVRQAPQGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPKGASVKVSKASGYTFTSYGISWVRQAPQGQGLEWMGWSAYNGNTNY 60

Qy 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVMTTDTSTAYMELSLRSDDTAVYYCARNPYTDSSEGFDFYWGQGLT 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180

Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

Query Match      81.2%; Score 1135; DB 52; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGIVTWVRQAPQGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPKGASVKVSKASGYTFTSYGISWVRQAPQGQGLEWMGWSAYNGNTNY 60

Qy 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVMTTDTSTAYMELSLRSDDTAVYYCARNPYTDSSEGFDFYWGQGLT 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180

Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

Query Match      81.2%; Score 1135; DB 52; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGIVTWVRQAPQGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPKGASVKVSKASGYTFTSYGISWVRQAPQGQGLEWMGWSAYNGNTNY 60

Qy 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVMTTDTSTAYMELSLRSDDTAVYYCARNPYTDSSEGFDFYWGQGLT 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180

Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253
```

RESULT 10

```
US-11-266-444-1964
; Sequence 1964, Application US/11266444
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523P1D1
; CURRENT APPLICATION NUMBER: US/11/266,444
; CURRENT FILING DATE: 2005-11-04
; PRIOR APPLICATION NUMBER: 09/880,746
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-266-444-1964
```

```
Query Match      81.2%; Score 1135; DB 42; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGIVTWVRQAPQGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPKGASVKVSKASGYTFTSYGISWVRQAPQGQGLEWMGWSAYNGNTNY 60

Qy 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVMTTDTSTAYMELSLRSDDTAVYYCARNPYTDSSEGFDFYWGQGLT 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180

Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253
```

```
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180
Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239
Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253
```

RESULT 11

```
US-60-725-626-1964
; Sequence 1964, Application US/60725626
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP8
; CURRENT APPLICATION NUMBER: US/60/725,626
; CURRENT FILING DATE: 2005-10-13
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-725-626-1964
```

```
Query Match      81.2%; Score 1135; DB 52; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
```

```
Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGIVTWVRQAPQGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPKGASVKVSKASGYTFTSYGISWVRQAPQGQGLEWMGWSAYNGNTNY 60

Qy 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVMTTDTSTAYMELSLRSDDTAVYYCARNPYTDSSEGFDFYWGQGLT 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180

Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253
```

RESULT 12

```
US-60-735-988-1964
; Sequence 1964, Application US/60735988
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP9
; CURRENT APPLICATION NUMBER: US/60/735,988
; CURRENT FILING DATE: 2005-11-14
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-735-988-1964
```

```
Query Match      81.2%; Score 1135; DB 52; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFSNYGIVTWVRQAPQGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKEPKGASVKVSKASGYTFTSYGISWVRQAPQGQGLEWMGWSAYNGNTNY 60

Qy 61 AQKFGQRTVMTTDTSTAYMELSLRSDDTAVYYCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVMTTDTSTAYMELSLRSDDTAVYYCARNPYTDSSEGFDFYWGQGLT 120

Qy 116 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGG--EIVLTQSPPLSLPVTLGQPASISCRSSQNLVYSDGNTYLSW 180

Qy 174 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
Db 181 FQRPQGQSPRLIYKVNRSRGVDPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239

Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253
```

```
Db 1 QVQLQQSGAEVKPGASVKVCKASGYTFSTYGISWVRQAPGQGLEWMGWISAYNGNTNY 60
Qy 61 AQKPGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDDE----PYWGQGTILVT 115
Db 61 AQKLGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDDE----PYWGQGTILVT 120
Qy 116 VSSGGGGGGGGGGGGGG--BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGGGG--BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 180
Qy 174 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTWHP 233
Db 181 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTWHP 239
Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

RESULT 13
US-60-776-665-1964
; Sequence 1964, Application US/60776665
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulat
; FILE REFERENCE: PF523PP10
; CURRENT APPLICATION NUMBER: US/60/776,665
; CURRENT FILING DATE: 2006-02-27
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-776-665-1964

Query Match 81.2%; Score 1135; DB 52; Length 253;
Best Local Similarity 86.2%; Pred. No. 1.8e-85;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;

Qy 3 QVQLVQSGAEVKPGASVKVCKASGYTFSTYGISWVRQAPGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQQSGAEVKPGASVKVCKASGYTFSTYGISWVRQAPGQGLEWMGWISAYNGNTNY 60
Qy 61 AQKPGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDDE----PYWGQGTILVT 115
Db 61 AQKLGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDDE----PYWGQGTILVT 120
Qy 116 VSSGGGGGGGGGGGGGG--BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 173
Db 121 VSSGGGGGGGGGGGGGG--BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSW 180
Qy 174 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTWHP 233
Db 181 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTWHP 239
Qy 234 PRTEGQGTKEIKR 247
Db 240 PFTFGQGTKEIKR 253

RESULT 14
US-09-202-000-14
; Sequence 14, Application US/09202000
; GENERAL INFORMATION:
; APPLICANT: Sawyer, Jaymie R.
; APPLICANT: Winter, Greg P.
; APPLICANT: Mahoney, Walter C.
; APPLICANT: Boehringer Mannheim Corp.
; TITLE OF INVENTION: COMPLEX SPECIFIC ANTIBODIES, METHOD OF
; FILE REFERENCE: 337462002200
; CURRENT APPLICATION NUMBER: US/09/202,000
```

```
; CURRENT FILING DATE: 1998-12-01
; EARLIER APPLICATION NUMBER: 60/067,428
; EARLIER FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PSA-ACT complex
US-09-202-000-14

Query Match 72.4%; Score 1011.5; DB 22; Length 242;
Best Local Similarity 78.7%; Pred. No. 3.1e-75;
Matches 196; Conservative 20; Mismatches 24; Indels 9; Gaps 3;

Qy 1 MAQVQLVQSGAEVKPGASVKVCKASGYTFSTYGISWVRQAPGQGLEWMGWIN--NGNT 58
Db 1 MAQVQLVQSGAEVKPGASVKVCKASGYTFSTYGISWVRQAPGQGLEWMGWISAYNGNT 60
Qy 59 HYAQKFGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDDEPYWGQGTILVTSS 118
Db 61 KYSQKFGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDPRGDDDEPYWGQGTILVTSS 115
Qy 119 GGGSGGGGGGGGGGGGG--BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOQRP 178
Db 116 GGGSGGGGGGGGGGGGG--BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFOQRP 175
Qy 179 GQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTWHPPTFG 238
Db 176 GQPPQLLIYEVNRFSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYCMQGTWHPPTFG 233
Qy 239 QGTKEIKR 247
Db 234 QGTKEIKR 242

RESULT 15
US-10-762-593-589
; Sequence 589, Application US/10762593
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Antibodies Against F1 and V Antigens of Yersinia Pestis
; FILE REFERENCE: PF603P1
; CURRENT APPLICATION NUMBER: US/10/762,593
; CURRENT FILING DATE: 2004-01-23
; NUMBER OF SEQ ID NOS: 1272
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 589
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: scFv protein YVG0277
US-10-762-593-589

Query Match 70.7%; Score 988.5; DB 37; Length 252;
Best Local Similarity 73.4%; Pred. No. 2.7e-73;
Matches 185; Conservative 35; Mismatches 25; Indels 7; Gaps 4;

Qy 3 QVQLVQSGAEVKPGASVKVCKASGYTFSTYGISWVRQAPGQGLEWMGWIN--NGNTHY 60
Db 1 EVQLLESGGGLVQPQGGSLRLSCAASGFTFSSYAMSVWRQAPGKLEWVSAISGGSGSTYY 60
Qy 61 AQKFGQGRVTMTDTSTAYMELSLRSDDTAVYYCARDP--RG-DDEPYWGQGTILVTSS 117
Db 61 ADSVKGRFTISRDNKNTLYLQWNSLRRAEDTAVYYCASDPPLEGYNLDYWGGRGTLVTSS 120
Qy 118 SGGSGGGGGGGGGGG--BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFO 175
Db 121 SGGSGGGGGGGGGGG--BIVLTQSPSLPVTILGQPASISCRSSQNLVYSDGNTYLSWFO 180
```


Qy	176	ORPGQSPRLIYKVSNRDSGVDPDRFSSGSGTDFTLKISRVEAEDIGVYYCMQGTWPPER	235
Db	181	ORPGQSPRLIYKVSNRDSGVDPDRFSSGSGTDFTLKISRVEAEDIGVYYCMQGTWPPER	240
Qy	236	TFGQGTKEIKR	247
Db	241	TFGQGTLEIKR	252

Search completed: August 9, 2006, 13:26:30
Job time : 605 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:16:37 ; Search time 34 Seconds
(without alignments)
511.336 Million cell updates/sec

Title: US-11-056-825-2
Perfect score: 1398
Sequence: 1 MAQVLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 294602 seqs, 66610880 residues

Total number of hits satisfying chosen parameters: 294602

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*

- 1: /EMC_Celerra_SID33/prodata/2/paa/PCT_NEW_COMB.pap.*
- 2: /EMC_Celerra_SID33/prodata/2/paa/US06_NEW_COMB.pap.*
- 3: /EMC_Celerra_SID33/prodata/2/paa/US07_NEW_COMB.pap.*
- 4: /EMC_Celerra_SID33/prodata/2/paa/US08_NEW_COMB.pap.*
- 5: /EMC_Celerra_SID33/prodata/2/paa/US09_NEW_COMB.pap.*
- 6: /EMC_Celerra_SID33/prodata/2/paa/US10_NEW_COMB.pap.*
- 7: /EMC_Celerra_SID33/prodata/2/paa/US11_NEW_COMB.pap.*
- 8: /EMC_Celerra_SID33/prodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1135	81.2	253	8	US-60-797-351-1964 Sequence 1964, Ap
2	1135	81.2	253	8	US-60-814-869-1964 Sequence 1964, Ap
3	1135	81.2	253	8	US-60-815-559-1964 Sequence 1964, Ap
4	1135	81.2	253	8	US-60-834-152-1964 Sequence 1964, Ap
5	924.5	66.1	244	8	US-60-797-351-1991 Sequence 1991, Ap
6	924.5	66.1	244	8	US-60-814-869-1991 Sequence 1991, Ap
7	924.5	66.1	244	8	US-60-815-559-1991 Sequence 1991, Ap
8	924.5	66.1	244	8	US-60-834-152-1991 Sequence 1991, Ap
9	919.5	65.8	277	7	US-11-177-495-3 Sequence 3, Appli
10	906	64.8	251	8	US-60-797-351-1238 Sequence 1238, Ap
11	906	64.8	251	8	US-60-814-869-1238 Sequence 1238, Ap
12	906	64.8	251	8	US-60-815-559-1238 Sequence 1238, Ap
13	906	64.8	251	8	US-60-834-152-1238 Sequence 1238, Ap
14	899	64.3	251	8	US-60-797-351-1921 Sequence 1921, Ap
15	899	64.3	251	8	US-60-814-869-1921 Sequence 1921, Ap
16	899	64.3	251	8	US-60-815-559-1921 Sequence 1921, Ap
17	899	64.3	251	8	US-60-834-152-1921 Sequence 1921, Ap
18	898.5	64.3	276	6	US-10-536-834A-95 Sequence 95, Appl
19	893.5	63.9	250	8	US-60-797-351-62 Sequence 62, Appl
20	893.5	63.9	250	8	US-60-814-869-62 Sequence 62, Appl
21	893.5	63.9	250	8	US-60-815-559-62 Sequence 62, Appl
22	893.5	63.9	250	8	US-60-834-152-62 Sequence 62, Appl
23	893	63.9	251	8	US-60-797-351-85 Sequence 85, Appl
24	893	63.9	251	8	US-60-814-869-85 Sequence 85, Appl
25	893	63.9	251	8	US-60-815-559-85 Sequence 85, Appl

26	893	63.9	251	8	US-60-834-152-85 Sequence 85, Appl
27	891.5	63.8	248	8	US-60-797-351-1446 Sequence 1446, Ap
28	891.5	63.8	248	8	US-60-814-869-1446 Sequence 1446, Ap
29	891.5	63.8	248	8	US-60-815-559-1446 Sequence 1446, Ap
30	891.5	63.8	248	8	US-60-834-152-1446 Sequence 1446, Ap
31	890.5	63.7	260	8	US-60-797-351-1039 Sequence 1039, Ap
32	890.5	63.7	260	8	US-60-814-869-1039 Sequence 1039, Ap
33	890.5	63.7	260	8	US-60-815-559-1039 Sequence 1039, Ap
34	890.5	63.7	260	8	US-60-834-152-1039 Sequence 1039, Ap
35	887	63.4	251	8	US-60-797-351-264 Sequence 264, App
36	887	63.4	251	8	US-60-814-869-264 Sequence 264, App
37	887	63.4	251	8	US-60-815-559-264 Sequence 264, App
38	887	63.4	251	8	US-60-834-152-264 Sequence 264, App
39	886	63.4	251	8	US-60-797-351-178 Sequence 178, App
40	886	63.4	251	8	US-60-814-869-178 Sequence 178, App
41	886	63.4	251	8	US-60-815-559-178 Sequence 178, App
42	886	63.4	251	8	US-60-834-152-178 Sequence 178, App
43	885	63.3	251	8	US-60-797-351-234 Sequence 234, App
44	885	63.3	251	8	US-60-814-869-234 Sequence 234, App
45	885	63.3	251	8	US-60-815-559-234 Sequence 234, App

ALIGNMENTS

RESULT 1
US-60-797-351-1964
; Sequence 1964, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator;
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1964

Query Match 81.2%; Score 1135; DB 8; Length 253;
Best Local Similarity 86.2%; Pred. No. 9.2e-81;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
QY 3 QVQLVQSGAEVKKPGASVKVCSGYSYTFNSGITWVROAPGQGLEWMGWIN--NGNTHY 60
Db 1 QVQLQSGAEVKKEPGASVKVCSGYSYTFNSGITWVROAPGQGLEWMGWIN--NGNTHY 60
QY 61 AQKFGQRTVTMTDSTSTAYMELRSLRSDDTAVYVCARDPRGDDE-----PYWGQGLT 115
Db 61 AQKLGQRTVTMTDSTSTAYMELRSLRSDDTAVYVCARNPYVYDSSEGFYWGQGLTWT 120
QY 116 VSSGGGGGGGGGGGGGG--EIVLTQSLPLSLPVTIGQPASISCRSSQNLVYSDGNLYLSW 173
Db 121 VSSGGGGGGGGGGGGGG--EIVLTQSLPLSLPVTIGQPASISCRSSQNLVYSDGNLYLSW 180
QY 174 FQORPGQSPRLIYKVNRDSCVPPRFGSGSGTFTLKISRVEAEDIGVYCMQGTHTP 233
Db 181 FQORPGQSPRLIYKVNRDSCVPPRFGSGSGTFTLKISRVEAEDIGVYCMQGTHTP 239
QY 234 PRTFGQGTKEIKR 247
Db 240 PRTFGQGTKEIKR 253

RESULT 2
US-60-814-869-1964
; Sequence 1964, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind to B Lymphocyte Stimulator;
; FILE REFERENCE: PF523PP12

; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-1964

Query Match 81.2%; Score 1135; DB 8; Length 253;
Best Local Similarity 86.2%; Pred. No. 9.2e-81;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
QY 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFTSYGISTWVRQAPQGQLEWMGWIN--NGNTHY 60
DB 1 QVQLQSGAEVKKPKGASVKVSKASGYTFTSYGISTWVRQAPQGQLEWMGWISAYNGNTNY 60
QY 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 115
DB 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 120
QY 116 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQAPASISCRSSQNLVYSDGNTYLSW 173
DB 121 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQAPASISCRSSQNLVYSDGNTYLSW 180
QY 174 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
DB 181 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239
QY 234 PRTEGQGTKEIKR 247
DB 240 PFTFGQGTKEIKR 253

RESULT 3
US-60-815-559-1964
; Sequence 1964, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-1964

Query Match 81.2%; Score 1135; DB 8; Length 253;
Best Local Similarity 86.2%; Pred. No. 9.2e-81;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
QY 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFTSYGISTWVRQAPQGQLEWMGWIN--NGNTHY 60
DB 1 QVQLQSGAEVKKPKGASVKVSKASGYTFTSYGISTWVRQAPQGQLEWMGWISAYNGNTNY 60
QY 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 115
DB 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 120
QY 116 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQAPASISCRSSQNLVYSDGNTYLSW 173
DB 121 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQAPASISCRSSQNLVYSDGNTYLSW 180
QY 174 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
DB 181 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239
QY 234 PRTEGQGTKEIKR 247
DB 240 PFTFGQGTKEIKR 253

RESULT 4
US-60-834-152-1964
; Sequence 1964, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1964
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1964

Query Match 81.2%; Score 1135; DB 8; Length 253;
Best Local Similarity 86.2%; Pred. No. 9.2e-81;
Matches 219; Conservative 15; Mismatches 10; Indels 10; Gaps 4;
QY 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFTSYGISTWVRQAPQGQLEWMGWIN--NGNTHY 60
DB 1 QVQLQSGAEVKKPKGASVKVSKASGYTFTSYGISTWVRQAPQGQLEWMGWISAYNGNTNY 60
QY 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 115
DB 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 120
QY 116 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQAPASISCRSSQNLVYSDGNTYLSW 173
DB 121 VSSGGSGGGGGGGGGG--EIVLTQSPPLSLPVTLGQAPASISCRSSQNLVYSDGNTYLSW 180
QY 174 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQGTHWP 233
DB 181 FQQRPGQSPRLIYKVNRSNDSGVPDRFSGSGSGTDFTLKISRVEAEDIGVYYCMQATRW- 239
QY 234 PRTEGQGTKEIKR 247
DB 240 PFTFGQGTKEIKR 253

RESULT 5
US-60-797-351-1991
; Sequence 1991, Application US/60797351
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP11
; CURRENT APPLICATION NUMBER: US/60/797,351
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-797-351-1991

Query Match 66.1%; Score 924.5; DB 8; Length 244;
Best Local Similarity 71.1%; Pred. No. 1.8e-64;
Matches 177; Conservative 34; Mismatches 29; Indels 9; Gaps 4;
QY 3 QVQLVQSGAEVKKPKGASVKVSKASGYTFTSYGISTWVRQAPQGQLEWMGWIN--NGNTH- -Y 60
DB 1 EYQLVQSGGGLVPPGSLRLSCAASGFTFTSSYMMNWVRQAPGKGLVWSSISSSSSHIY 60
QY 61 AQKQGRVTMTTDTSTSTAYMELSLRSDDTAVYYCARDPRGDDP-----PYWGQGLT 120
DB 61 ADSVKGRTISRDNKNSLYLQWNSLRADTAIVYCARVHSSGS---WGQGLTVTVSSGG 117
QY 121 GSGSGSGGGGGG--EIVLTQSPPLSLPVTLGQAPASISCRSSQNLVYSDGNTYLSWFOORP 178

Db 118 GSGGGGGGGGSGGALDVVMTQSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 177
QY 179 QGSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 238
Db 178 QGSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 235
QY 239 QGKVEIKR 247
Db 236 QGTRLEIKR 244

RESULT 6
US-60-814-869-1991
; Sequence 1991, Application US/60814869
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP12
; CURRENT APPLICATION NUMBER: US/60/814,869
; CURRENT FILING DATE: 2006-06-20
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-814-869-1991

Query Match 66.1%; Score 924.5; DB 8; Length 244;
Best Local Similarity 71.1%; Pred. No. 1.8e-64;
Matches 177; Conservative 34; Mismatches 29; Indels 9; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRQAPGQGLEWMGWINNNGTH--Y 60
Db 1 EVQLVQSGGGLVKPGGSLRLSCAASGFTFSSYMNWVRQAPGKGLVWVSSISSSSHIYY 60
QY 61 AQKFGQRTVMTDTSTSTAYMELSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 ADSVKGRFTISRDNKNSLYLQMSLRADTAIVYICARVHSSGS---WGQGLTVTVSSGG 117
QY 121 GSGGGGGGGGSGG--EIVLTQSPSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 178
Db 118 GSGGGGGGGGSGGALDVVMTQSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 177
QY 179 QGSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 238
Db 178 QGSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 235
QY 239 QGKVEIKR 247
Db 236 QGTRLEIKR 244

RESULT 7
US-60-815-559-1991
; Sequence 1991, Application US/60815559
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP13
; CURRENT APPLICATION NUMBER: US/60/815,559
; CURRENT FILING DATE: 2006-06-22
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-815-559-1991

Query Match 66.1%; Score 924.5; DB 8; Length 244;
Best Local Similarity 71.1%; Pred. No. 1.8e-64;
Matches 177; Conservative 34; Mismatches 29; Indels 9; Gaps 4;
QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRQAPGQGLEWMGWINNNGTH--Y 60

Db 1 EVQLVQSGGGLVKPGGSLRLSCAASGFTFSSYMNWVRQAPGKGLVWVSSISSSSHIYY 60
QY 61 AQKFGQRTVMTDTSTSTAYMELSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 ADSVKGRFTISRDNKNSLYLQMSLRADTAIVYICARVHSSGS---WGQGLTVTVSSGG 117
QY 121 GSGGGGGGGGSGG--EIVLTQSPSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 178
Db 118 GSGGGGGGGGSGGALDVVMTQSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 177
QY 179 QGSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 238
Db 178 QGSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 235
QY 239 QGKVEIKR 247
Db 236 QGTRLEIKR 244

RESULT 8
US-60-834-152-1991
; Sequence 1991, Application US/60834152
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
; FILE REFERENCE: PF523PP14
; CURRENT APPLICATION NUMBER: US/60/834,152
; CURRENT FILING DATE: 2006-07-31
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1991
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-834-152-1991

Query Match 66.1%; Score 924.5; DB 8; Length 244;
Best Local Similarity 71.1%; Pred. No. 1.8e-64;
Matches 177; Conservative 34; Mismatches 29; Indels 9; Gaps 4;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITVWRQAPGQGLEWMGWINNNGTH--Y 60
Db 1 EVQLVQSGGGLVKPGGSLRLSCAASGFTFSSYMNWVRQAPGKGLVWVSSISSSSHIYY 60
QY 61 AQKFGQRTVMTDTSTSTAYMELSLRSDDTAVYICARDPRGDDPEYWGQGLTVTVSSGG 120
Db 61 ADSVKGRFTISRDNKNSLYLQMSLRADTAIVYICARVHSSGS---WGQGLTVTVSSGG 117
QY 121 GSGGGGGGGGSGG--EIVLTQSPSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 178
Db 118 GSGGGGGGGGSGGALDVVMTQSLPVLTPVTLGQPASISCRSSQSLVHSDGNTYLNWFOQRP 177
QY 179 QGSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 238
Db 178 QGSPRLIYKVSNRDGVDPDRFSGSGSGTDTFTLKISRVEAEDIGVYCMQGTHTVTPRTFG 235
QY 239 QGKVEIKR 247
Db 236 QGTRLEIKR 244

RESULT 9
US-11-177-495-3
; Sequence 3, Application US/11177495
; GENERAL INFORMATION:
; APPLICANT: RAPAPORT FAMILY INSTITUTE FOR RESEARCH
; TITLE OF INVENTION: METHODS OF DETECTING HAPTOGLOBIN PHENOTYPE
; FILE REFERENCE: P-7018-US
; CURRENT APPLICATION NUMBER: US/11/177,495
; CURRENT FILING DATE: 2005-07-11
; PRIOR APPLICATION NUMBER: USP 60/586,733
; PRIOR FILING DATE: 2004-07-12
; NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn version 3.3

SEQ ID NO 3
LENGTH: 277
TYPE: PRT
ORGANISM: Human
US-11-177-495-3

Query Match 65.8%; Score 919.5; DB 7; Length 277;
Best Local Similarity 69.4%; Pred. No. 4.9e-64;
Matches 175; Conservative 39; Mismatches 29; Indels 9; Gaps 4;

QY 1 MAQVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGW--NGNT 58
DB 1 MAEVQSQSGTELVKPGASVKLSCKASGYTFSNYGITWVRQAPGQGLEWMGIHHPYSGL 60
QY 59 HYAQFGQVRVTMTDTSTAYMELSLRSDDTAVYICARDPRGDDE---PYWGGTGLVT 115
DB 61 KFNFRFNKATLTVDKSSNTAYMQVSSLTSEDSAVYICARD---DDSAWFAYWGGTGLVT 117
QY 116 VSSGGGSGGGGGGGGSEIVLTQSPISLPTVLTGQPASISCRSSQNLVYSDGNTYLSWFQ 175
DB 118 VSSGGGSGGGGGGGGSDVVMTQTPLSLPVSLGDPQASISCRSSQSIVHSGNTYLEWYL 177
QY 176 QPQSPRLIYKVNDRSGVDPDRFSGSGGTDFTLKISRVEADIGVYCMQGTHTWPPR 235
DB 178 QKQSPKLLIYKVNDRSGVDPDRFSGSGGTDFTLKISRVEADIGVYCFQGSHPVW 236
QY 236 TFGGQTKVEIKR 247
DB 237 TFGGQTKLEIKR 248

RESULT 10

US-60-797-351-1238
Sequence 1238, Application US/60797351
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
FILE REFERENCE: PF523PP11
CURRENT APPLICATION NUMBER: US/60/797,351
CURRENT FILING DATE: 2006-05-04
NUMBER OF SEQ ID NOS: 3247
SEQ ID NO 1238
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-60-797-351-1238

Query Match 64.8%; Score 906; DB 8; Length 251;
Best Local Similarity 67.7%; Pred. No. 4.9e-63;
Matches 178; Conservative 29; Mismatches 26; Indels 30; Gaps 5;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGW--INNNGNTHY 60
DB 1 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWSAHDNDTKY 60
QY 61 AQKFGQVRVTMTDTSTAYMELSLRSDDTAVYICARDPRGDDEPY----- 107
DB 61 AQKFGQVRVTMTDTSTAYMELSLRSDDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 113
QY 108 -WGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPTVLTGQPASISCRSSQNLVY 164
DB 114 NWGQGLTVTVSSGGGGGGGGGGGSALETTLTQSPATLSVSPGERVTLSCRASQSV-- 171
QY 165 SDGNTYLSWFOQRPGQSPRLIYKVNDRSGVDPDRFSGSGGTDFTLKISRVEADIGVY 224
DB 172 ---SSNLAWFOQKPGQAPRLIYGASTRFTGIPARFSGSGGTDFSLTISLQSEDPAVY 228
QY 225 YCMQGTHTWPPRTFGQTKVEIKR 247
DB 229 YCQYNNWNPPIITFGQGTLEIKR 251

RESULT 11

US-60-814-869-1238
Sequence 1238, Application US/60814869
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
FILE REFERENCE: PF523PP12
CURRENT APPLICATION NUMBER: US/60/814,869
CURRENT FILING DATE: 2006-06-20
NUMBER OF SEQ ID NOS: 3247
SEQ ID NO 1238
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-60-814-869-1238

Query Match 64.8%; Score 906; DB 8; Length 251;
Best Local Similarity 67.7%; Pred. No. 4.9e-63;
Matches 178; Conservative 29; Mismatches 26; Indels 30; Gaps 5;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGW--INNNGNTHY 60
DB 1 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWSAHDNDTKY 60
QY 61 AQKFGQVRVTMTDTSTAYMELSLRSDDTAVYICARDPRGDDEPY----- 107
DB 61 AQKFGQVRVTMTDTSTAYMELSLRSDDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 113
QY 108 -WGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPTVLTGQPASISCRSSQNLVY 164
DB 114 NWGQGLTVTVSSGGGGGGGGGGGSALETTLTQSPATLSVSPGERVTLSCRASQSV-- 171
QY 165 SDGNTYLSWFOQRPGQSPRLIYKVNDRSGVDPDRFSGSGGTDFTLKISRVEADIGVY 224
DB 172 ---SSNLAWFOQKPGQAPRLIYGASTRFTGIPARFSGSGGTDFSLTISLQSEDPAVY 228
QY 225 YCMQGTHTWPPRTFGQTKVEIKR 247
DB 229 YCQYNNWNPPIITFGQGTLEIKR 251

RESULT 12

US-60-815-559-1238
Sequence 1238, Application US/60815559
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunospecifically Bind to B Lymphocyte Stimulator
FILE REFERENCE: PF523PP13
CURRENT APPLICATION NUMBER: US/60/815,559
CURRENT FILING DATE: 2006-06-22
NUMBER OF SEQ ID NOS: 3247
SEQ ID NO 1238
LENGTH: 251
TYPE: PRT
ORGANISM: Homo sapiens
US-60-815-559-1238

Query Match 64.8%; Score 906; DB 8; Length 251;
Best Local Similarity 67.7%; Pred. No. 4.9e-63;
Matches 178; Conservative 29; Mismatches 26; Indels 30; Gaps 5;

QY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGW--INNNGNTHY 60
DB 1 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQGLEWMGWSAHDNDTKY 60
QY 61 AQKFGQVRVTMTDTSTAYMELSLRSDDTAVYICARDPRGDDEPY----- 107
DB 61 AQKFGQVRVTMTDTSTAYMELSLRSDDTAVYICARDPRGDDEPY-----PYDPLTAYTFQYFG 113
QY 108 -WGQGLTVTVSSGGGGGGGGGGGSEIVLTQSPISLPTVLTGQPASISCRSSQNLVY 164
DB 114 NWGQGLTVTVSSGGGGGGGGGGGSALETTLTQSPATLSVSPGERVTLSCRASQSV-- 171

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:09:51 ; Search time 40 Seconds
(without alignment)
627.814 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKKPGASV.....KVEIKRGLGLVDYKDDDDK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_80.*

1: Pirl.*

2: Pirl.*

3: Pirl.*

4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	808	57.8	249	2 S41374	single chain Fv antibody
2	676.5	48.4	268	2 A56446	Ig heavy chain V r
3	660	47.2	233	2 J053322	p53 specific singl
4	565	40.4	114	2 B49002	Ig kappa chain V r
5	562	40.2	114	2 S49572	Ig kappa chain pre
6	561.5	40.2	133	2 S23230	Ig kappa chain pre
7	560.5	40.1	133	1 K2HURP	Ig kappa chain pre
8	553	39.6	140	2 S22658	Ig kappa chain pre
9	552.5	39.5	133	1 A24452	Ig kappa chain pre
10	551.5	39.4	133	2 S40324	Ig kappa chain V r
11	550.5	39.4	122	2 S40338	Ig kappa chain - h
12	547.5	39.2	142	2 S22902	Ig kappa chain V r
13	546.5	39.1	133	2 S42611	HUNVK protein prec
14	535.5	38.3	118	2 S40374	Ig kappa chain - h
15	529.5	37.9	132	2 S40322	Ig kappa chain - h
16	521.5	37.3	160	2 P0105	anti-PR2 erythrocy
17	521	37.3	120	2 S42268	Ig kappa chain V r
18	513	36.7	120	2 S42267	Ig kappa chain V r
19	509.5	36.4	122	2 S36271	Ig heavy chain V r
20	507	36.3	134	2 S40376	Ig kappa chain - h
21	505.5	36.2	129	2 S36260	Ig heavy chain V r
22	501.5	35.9	126	2 S40312	Ig kappa chain - h
23	496	35.5	124	2 S19665	Ig heavy chain V r
24	487	34.8	118	2 S36265	Ig heavy chain V r
25	486	34.8	135	2 S49530	anti-Sm antibody V
26	478.5	34.2	142	2 A32483	Ig heavy chain V r
27	473.5	33.9	125	2 S68170	Ig heavy chain V r
28	471.5	33.7	128	2 S40373	Ig kappa chain - h
29	471.5	33.7	219	2 S52028	Ig kappa chain - m

RESULT 1

S41374

single chain Fv antibody - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995

C:Accession: S41374

R:Artsaenko, O.; Weiler, E.W.; Muentz, K.; Conrad, U.

submitted to the EMBL Data Library, January 1994

A:Description: Construction and functional characterization of a single chain Fv antibody

A:Reference number: S41374

A:Accession: S41374

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-249 <ART>

A:Cross-references: UNIPARC:UPI000017C74A; EMBL:Z29480

Query Match

Best Local Similarity 57.8%; Score 808; DB 2; Length 249;

Matches 161; Conservative 33; Mismatches 50; Indels 4; Gaps 3;

QY	3	QVQLVQSGAEVKKPGASVKVKCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTHY	60
DB	1	QVQLVQSGAEVKKPGASVKLSCTASGFNPKDYYHWVKQRPKEGLEWIAIPASGNVKY	60
QY	61	AQKFGQRTVMTDTSTAYMELRLSRSDTAVVYCA-RDPRGDDEPYWGQGLTVTVSSG	119
DB	61	VPRFDQKATITADTSNTAYLLLSLTSEDVAVVYCARDTLYTSLGYWGQGSTVTVSSR	120
QY	120	GGSGGGSGGGGGSHVLTQSPFLSLPVTILGPASISCRSSQNLVYSDGNTYLSWFOQRP	179
DB	121	GGSGGGSGGGGGSDIELTQSPFVVIPEGSVISCRSSKSLYSDGSLFWFLQRP	180
QY	180	QSPRLIYKVNDRDGVDPDFSGSGSGTDFTLKISRVEAEDIGVYCMQTHWPPFTFGQ	239
DB	181	QSPQLLIYRMSNLASGVDPDFSGSGSGTFTLKISRVEAEDIGVYCMQTHREY-PLTFGA	239
QY	240	GTKVEIKR	247
DB	240	GTKLEIKR	247

RESULT 2

A56446

Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)

C:Species: Mus musculus (house mouse)

C:Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 16-Aug-1996

C:Accession: A56446

R:Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.

J. Biol. Chem. 270, 7829-7835, 1995

A:Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical

A:Reference number: A56446; MUID:95229583; PMID:7713873

A:Accession: A56446

A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-268 <TAN>
A:Cross-references: UNIPARC:UPI000017C6D0; GB:U20617
C;Keywords: heterotetramer; immunoglobulin

Query Match 48.4%; Score 676.5; DB 2; Length 268;
Best Local Similarity 54.6%; Pred. No. 1.9e-40;
Matches 136; Conservative 36; Mismatches 68; Indels 9; Gaps 3;

Qy 1 MAQVLVQSGAEVKKPGASVKSCASGVTSNYGITWVROAPGQGLEWMGWI--NNGNT 58
Db 1 MAQVKLQSGAEVLKPGASVKLCTSTGTFNIDKIDYTHHWKQPQGLRWIGRIAPANGIT 60

Qy 59 HYAKQFGQGVTTTDTSTSTAYMELRSDDTAVYICARDPRGDDPEYWGQGTILVTVSS 118
Db 61 KYDPKFGQKATIAADTSNTAYLQLSLTSEDVAVYICASYLYTRYENYWGQGTIVTVSS 120

Qy 119 GGGSGGGGGGGGSEIVLTQSPSLPVLTLGPASISCRSSQNLVYSDGNTYLSWFOQRP 178
Db 121 GGGSGGGGDSGGSDIELTQSPALMSASLGEKVTMSCRASSV-----NEIYVYQOKS 174

Qy 179 GQSPRLIYKVNRSQVDPDPFGSGSGCTDFTLKISRVEAEDIGVYICMQGTHWPPRTFG 238
Db 175 DASPKLWYTYTSHLPVGPVAPRFGSGSGNSYSLTISMEGEDAATYICQOFTS--SPFTFG 233

Qy 239 QGTRKVEIKR 247
Db 234 SGTKLEIKR 242

RESULT 3
JC5322
p53 specific single-chain antibody Pab421 - human
C;Species: Homo sapiens (man)
C;Date: 15-May-1997 #sequence_revision 15-May-1997 #text_change 18-Jul-1997
C;Accession: JC5322
R;Jannot, C.B.; Hynes, N.E.
Biochem. Biophys. Res. Commun. 230, 242-246, 1997
A;Title: Characterization of scFv-421, a single-chain antibody targeted to p53.
A;Reference number: JC5322; MUID:97168950; PMID:9016757
A;Accession: JC5322
A:Molecule type: mRNA
A:Residues: 1-233 <JAN>
A:Cross-references: UNIPARC:UPI000017C2DE
A;Experimental source: hybridoma cell
C;Comment: This protein specifically binds the tumor suppressor protein p53. It restores

Query Match 47.2%; Score 660; DB 2; Length 233;
Best Local Similarity 58.3%; Pred. No. 2.3e-39;
Matches 140; Conservative 33; Mismatches 57; Indels 10; Gaps 6;

Qy 8 QSGAEVKKPGASVKVSCASGVTSNYGITWVROAPGQGLEWMGIN--NGNTHYAQKFP 65
Db 2 ESGAEVRSASVKLSCTTSGTFNIDYTHHWKQPEQGLEWIGRIDPENGADMTTRSGG 61

Qy 66 GRVTMTTSTSTAYMELRSDDTAVYICARDPRGDDPEYWGQGTILVTVSSGGGGSGG 125
Db 62 VKATMTADTSNTAYLQLSLTSEDVAVYIC---NAGMD--YWGQGTIVTVSSGGGGSGG 116

Qy 126 GSGGGGSEIVLTQSPSLPVLTLGPASISCRSSQNLVYSDGNTYLSWFOQRPQGSPLRL 185
Db 117 RASGGGSDIELTQSPASLAVSLGQRTATISCRASKS--VSTSGSYTHWNQKQFGPPRL 175

Qy 186 IYKVNRSQVDPDRFGSGSGCTDFTLKISRVEAEDIGVYICMQGTHWPPRTFGQTKVEI 245
Db 176 IYLVNLSGVPARFGSGSGCTDFTLNHPVEEEDAATYIC-QHIRELTRSEG-GTKLEI 233

RESULT 4
B49002
Ig kappa chain V region, rheumatoid factor RF antibody - human (fragment)
C;Species: Homo sapiens (man)

C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B49002
R;Stuber, F.; Lee, S.K.; Bridges Jr., S.L.; Koopman, W.J.; Schroeder, H.W.J.; Gaskin, F.;
Arthritis Rheum. 35, 900-904, 1992
A;Title: A rheumatoid factor from a normal individual encoded by VH2 and V kappa II gene
A;Reference number: A49002; MUID:92352481; PMID:1322670
A;Accession: B49002
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-114 <STU>
A:Cross-references: UNIPROT:Q9UL80; UNIPARC:UPI0000176B43
A;Experimental source: EBV-transformed lymphoblastoid cell line SSH23
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:110263, NCBIP:110264)
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;16-95/Domain: immunoglobulin homology <IMM>

Query Match 40.4%; Score 565; DB 2; Length 114;
Best Local Similarity 92.1%; Pred. No. 4.9e-33;
Matches 105; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 134 EIVLTQSPSLPVLTLGPASISCRSSQNLVYSDGNTYLSWFOQRPQGSPLRLIYKVSNRD 193
Db 1 DVVMTQSPSLPVLTLGPASISCRSSQSLVYSDGNTSLAWFOQRPQGSPLRLIYKVSNRD 60

Qy 194 SGVPDRFGSGSGTDFTLKISRVEAEDIGVYICMQGTHWPPRTFGQTKVEIKR 247
Db 61 SGVPDRFGSGSGTDFTLKISRVEAEDVGYYICMQGTHWPPPLTFGGGTKVEIKR 114

RESULT 5
S49572
Ig kappa chain precursor - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 21-Jan-2000
C;Accession: S49572
R;Giachino, C.; Padova, E.; Lanzavecchia, A.
submitted to the EMBL Data Library, November 1994
A;Description: k+l+ dual receptor B cells are present in the human peripheral repertoire.
A;Reference number: S49571
A;Accession: S49572
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-114 <GIA>
A:Cross-references: UNIPARC:UPI0000116709; EMBL:Z46626; NID:g575261; PIDN:CAA86596.1; PII
C;Superfamily: immunoglobulin V region; immunoglobulin homology
F;16-95/Domain: immunoglobulin homology <IMM>

Query Match 40.2%; Score 562; DB 2; Length 114;
Best Local Similarity 89.5%; Pred. No. 7.9e-33;
Matches 102; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Qy 134 EIVLTQSPSLPVLTLGPASISCRSSQNLVYSDGNTYLSWFOQRPQGSPLRLIYKVSNRD 193
Db 1 DVVMTQSPSLPVLTLGPASISCRSSQSLVYDGTSLAWFOQRPQGSPLRLIYKVSNRD 60

Qy 194 SGVPDRFGSGSGTDFTLKISRVEAEDIGVYICMQGTHWPPRTFGQTKVEIKR 247
Db 61 SGVPDRFGSGSGTDFTLKISRVEAEDVGYYICIQGTHWPPQYTFGGGTKLEIKR 114

RESULT 6
S23230
Ig kappa chain precursor V-J region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C;Accession: S23230
R;Kennedy, M.A.
J. Exp. Med. 173, 1033-1036, 1991
A;Title: Novel chromosome translocation caused by fusion of immunoglobulin heavy and light
A;Reference number: S23230; MUID:91178438; PMID:1840606
A;Accession: S23230

S2258
Ig kappa chain precursor V region (0-81VL) - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 31-Dec-2004
C/Accession: S22658
R/Hirabayashi, Y.; Munakata, Y.; Sasaki, T.; Sano, H.
Nucleic Acids Res. 20, 2601, 1992
A/Title: Variable regions of a human anti-DNA antibody O-81 possessing lupus nephritis-associated disulfide bonds.
A/Reference number: S22657; MUID:92285150; PMID:1598223
A/Accession: S22658
A/Molecule type: mRNA
A/Residues: 1-140 <HR>
A/Cross-references: UNIPROT:Q8TCD0; UNIPARC:UPI00001769CF; EMBL:X59135
C/Superfamily: immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-140/Product: Ig kappa chain (fragment) #status predicted <MAT>
F/36-115/Domain: immunoglobulin homology <IMM>

Query Match 39.6%; Score 553; DB 2; Length 140;
Best Local Similarity 85.6%; Pred. No. 4.2e-32;
Matches 101; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 130 GGGSEIVLTOSPLSLPVTLGOPASISCRSSONLVSDGNTYLSWFOORPGQSRRLIYKV 189
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
17 GSGDGVVMTQSP LSLPVT LGOPAS ISCRSSQLVSHDSGNTYLNWFWOORPGQSPRLIYRV 76

QY 190 SNRDGVPDRFGSGSGGTDFTLKISRVEADIGVYCMQGTHTWPPRTFGQGTKVEIKR 247
Db |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
77 SNRDGVPDRFGSGSGGTDFTLKISRVEADIGVYCMQGTHTWSPITFGQGTREIKR 134

RESULT 9
A24452
Ig kappa chain precursor V-II region (RPMI 6410) - human
C/Species: Homo sapiens (man)
C/Date: 24-Jan-1988 #sequence_revision 09-Aug-1996 #text_change 16-Jul-1999
C/Accession: A24452
R/Weir, L.; Leder, P.
Nucleic Acids Res. 14, 3957-3970, 1986
A/Title: Structure and expression of a human subgroup II immunoglobulin kappa gene.
A/Reference number: A24452; MUID:86232631; PMID:3086847
A/Accession: A24452
A/Molecule type: DNA
A/Residues: 1-133 <MBI>
A/Cross-references: UNIPARC:UPI0000113B46; GB:M36859; NID:g185932; PID:AAAS58920.1; PID:AAAS58920.2
A/Note: this sequence was determined from the differentiated gene
C/Genetics:
A/Gene: GDB:IGKV2
A/Cross-references: GDB:I36265
A/Map position: 2p12-2p12
A/Introns: 17/1
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGW, the subunits associate into larger oligomers.
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-133/Product: Ig kappa chain V-II region (RPMI 6410) #status predicted <MAT>
F/36-115/Domain: immunoglobulin homology <IMM>
F/43-113/Disulfide bonds: #status predicted

Query Match 39.5%; Score 552.5; DB 1; Length 133;
Best Local Similarity 89.0%; Pred. No. 4.3e-32;
Matches 105; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

QY 130 GGGSEIVLTOSPLSLPVTLGOPASISCRSSONLVSDGNTYLSWFOORPGQSRRLIYKV 189
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
17 GSGDGVVMTQSP LSLPVT LGOPAS ISCRSSQLVSHDSGNTYLNWFWOORPGQSPRLIYKV 76

QY 190 SNRDGVPDRFGSGSGGTDFTLKISRVEADIGVYCMQGTHTWPPRTFGQGTKVEIKR 247
Db |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
77 SNRDGVPDRFGSGSGGTDFTLKISRVEADIGVYCMQGTHTW-SWTFQGQTKVEIKR 133

Search completed: August 9, 2006, 13:15:24
Job time : 41 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2006, 13:06:26 ; Search time 300 Seconds
(without alignments)
804.764 Million cell updates/sec

Title: US-11-056-825-2

Perfect score: 1398

Sequence: 1 MAQVQLVQSGAEVKPKPGASV.....KVEIKRGLGLVDYKDDDK 261

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	848.5	60.7	244	2	Q65ZC8_HUMAN
2	828	59.2	248	2	Q65ZQ7_MOUSE
3	764.5	54.7	255	2	Q6KB05_MOUSE
4	730.5	52.3	240	2	Q65ZC9_HUMAN
5	721.5	51.6	243	2	Q7TQM2_MOUSE
6	718.5	51.4	241	2	Q921A6_MOUSE
7	717.5	51.3	487	2	Q925S1_MOUSE
8	664.5	47.5	218	2	Q925S1_MOUSE
9	565	40.4	114	2	Q9UL80_HUMAN
10	560.5	40.1	133	1	KV2F_HUMAN
11	557.5	39.9	239	2	Q8TCD0_HUMAN
12	493.5	35.3	239	2	Q6P491_HUMAN
13	493.5	35.3	500	2	Q6N091_HUMAN
14	477.5	34.2	469	2	Q7Z7P5_HUMAN
15	471.5	33.7	518	2	Q6N030_HUMAN
16	471	33.7	115	2	Q5F210_MOUSE
17	465.5	33.3	239	2	Q8NEK0_HUMAN
18	462	33.0	124	2	Q9UL92_HUMAN
19	461.5	33.0	125	2	Q9UL95_HUMAN
20	458.5	32.8	113	1	KV2G_MOUSE
21	458.5	32.8	119	2	Q9UL94_HUMAN
22	456.5	32.7	117	1	KV2E_HUMAN
23	455.5	32.6	239	2	Q58E08_MOUSE
24	451	32.3	240	2	Q6PIH6_HUMAN
25	450.5	32.2	113	1	KV2B_HUMAN
26	450	32.2	498	2	Q6N041_HUMAN
27	448.5	32.1	113	1	KV2D_HUMAN
28	446	31.9	170	2	Q925S2_MOUSE
29	445	31.8	500	2	Q9BRV0_HUMAN
30	444	31.8	159	2	Q96Q80_HUMAN
31	438.5	31.4	219	2	Q65ZC0_MOUSE

32	433.5	31.0	480	2	Q6P089_HUMAN
33	428	30.6	86	2	Q7Z3V5_HUMAN
34	426	30.5	115	1	KV2A_HUMAN
35	425.5	30.4	119	2	Q9GYZ2_MOUSE
36	425	30.4	147	1	HVIC_HUMAN
37	422	30.2	112	1	KV2C_HUMAN
38	421	30.1	117	1	HVIB_HUMAN
39	419.5	30.0	458	2	Q5BJZ2_RAT
40	417.5	29.9	112	2	Q53VP8_MOUSE
41	417	29.8	116	2	Q9UL89_HUMAN
42	415.5	29.7	475	2	Q6N095_HUMAN
43	414.5	29.6	497	2	Q8WY24_HUMAN
44	414	29.6	519	2	Q5EBM2_HUMAN
45	410.5	29.4	234	2	Q5XKG4_MOUSE

ALIGNMENTS

RESULT 1
Q65ZC8_HUMAN PRELIMINARY; PRT; 244 AA.
AC Q65ZC8; integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Single-chain Fv (Fragment).
GN Name=scFv;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97362799; PubMed=9219263; DOI=10.1038/nbr0797-629;
RA Kontermann R.E., Wing M.G., Winter G.;
RT "Complement recruitment using bispecific diabodies";
RL Nat. Biotechnol. 15:629-631(1997).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; Y13057; CAA73500.1; -; mRNA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IG; 2.
DR PROSITE; PS50835; IG LIKE; 2.
KW Immunoglobulin domain.
FT NON_TER 1
FT NON_TER 244
SQ SEQUENCE 244 AA; 26127 MW; 4B1F17868338F28F CRC64;

Query Match	60.7%;	Score	848.5;	DB	2;	Length	244;
Best Local Similarity	65.6%;	Pred. No.	6.1e-61;				
Matches	164;	Conservative	34;	Mismatches	41;	Indels	11; Gaps 4;
QY	3	QVQLVQSGAEVKPKPGASVKVCKASGYTFSNYGITWVRQAPGQGLEWMGWI--NNGNTHY	60				
Db	1	QVQLVQSGAEVKPKGDSVKVCKASGYTFSDHYMHVVRQAPGQGLEWMGWIDPNNGDTRF	60				
QY	61	AQKPGFVRVTMTDTSTAYMELRLSRDSTAVYYCARDPRGD---DEPTWGQGLTVTVS	117				
Db	61	AQKPGFVRVTMTDTSTAYMELRLSRDSTAVYYCARDPRGD---DEPTWGQGLTVTVS	120				
QY	118	SGGGSGGGSGGGSGEIVLTQSLPVLTPGAPASISCRSSQNLVYSDGNVTYLSWFOQR	177				
Db	121	SGGGSGGGSGGGSGGGSDIQMTQSPSLASIGDRVTITCRASEGIYH----WLAWTQQK	175				
QY	178	PQSPRLIYKVSNRDSDGVPDRFSGSGSGTDFTLKISRVEAEDIGVYVCMQGTWPPRTF	237				

```
Db 176 PGKAPKFLIYKASLASGAPRFGSGSGTDTLTITSSLOPDDFATYYCQOYSNY-PLTF 234
Qy 238 GQGTKEIKR 247
Db 235 GGGTKLEIKR 244

RESULT 2
Q65ZQ7_9MURI
ID Q65ZQ7_9MURI PRELIMINARY; PRT; 248 AA.
AC Q65ZQ7;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE B3(Fv)-PE40 (Fragment).
GN Name=B3(Fv)-PE40;
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92020904; PubMed=1924323;
RA Brinkmann U., Pai L.H., FitzGerald D.J., Willingham M., Pastan I.;
RT "B3(Fv)-PE38DEL, a single-chain immunotoxin that causes complete
RT regression of a human carcinoma in mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8616-8620(1991).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; S57590; AAB19571.2; -; mRNA.
DR SMR; Q65ZQ7; 4-247.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 2.
KW Immunoglobulin domain.
FT NON_TER 248 248
SQ SEQUENCE 248 AA; 26634 MW; 7A3759B43B570950 CRC64;

Query Match 59.2%; Score 828; DB 2; Length 248;
Best Local Similarity 64.2%; Pred. No. 2.9e-59;
Matches 158; Conservative 41; Mismatches 43; Indels 4; Gaps 3;

Qy 4 VOLVQSGAEVKKPGASVKASCKGYSNYGIVTWRQAPGGLEWMGWINNTH--YA 61
Db 3 VKLVESGGGLVQPGGSLKLSKTSATGFTFSDYYWYWRQPEKRLSEWYVINSDDSAAYS 62
Qy 62 QKQGRVMTTDTSTSTAYMELRSDDTAVYYCARD--PRGDDDPYWGQGLTVTVSSGG 120
Db 63 DTVKGRFTISRDNKNTLYLQMSRLKSEDTAIVSCARGLANGAWFAYWGQGLTVTVSSGG 122
Qy 121 GSGGGGGGGGGSEIVLTQSPSLPVTLPGLPQASISCRSSQNLVYSDGNTLYSNFQORPGQ 180
Db 123 GSGGGGGGGGGSDVLTQSPSLPVLSDQASISCRSSQIIIVHSNGNTLYEWYLQKPGQ 182
Qy 181 SPRLIYKVNRSQVPRDFSGSGSGTDTLTLSRVEAEIDGVYCMQGTWHPPTFGQ 240
Db 183 SPKLIYKVNRSFSGVPRDFSGSGSGTDTLTLSRVEAEIDLVGYCFQGSF--VPFTFGS 241
Qy 241 TKVEIK 246
Db 242 TKLEIK 247

RESULT 3
Q6KB05_MOUSE
ID Q6KB05_MOUSE PRELIMINARY; PRT; 255 AA.
AC Q6KB05;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE scFv B8E5 protein (Fragment).
GN Name=scFv B8E5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=Balb/c;
RX PubMed=15485827; DOI=10.1074/jbc.M407213200;
RA Peter J.C., Wallukat G., Tugler J., Maurice D., Roegel J.C.,
RA Briand J.P., Hoebeke J.;
RT "Modulation of the M2 muscarinic acetylcholine receptor activity with
RT monoclonal anti-M2 receptor antibody fragments.";
RL J. Biol. Chem. 279:55697-55706(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AJ746180; CAG34081.1; -; Other_DNA.
DR HSSP; P01837; 1KCR.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 2.
KW Immunoglobulin domain.
FT NON_TER 1 1
SQ SEQUENCE 255 AA; 27445 MW; B68BD38395DF713B CRC64;

Query Match 54.7%; Score 764.5; DB 2; Length 255;
Best Local Similarity 59.6%; Pred. No. 4.5e-54;
Matches 149; Conservative 42; Mismatches 52; Indels 7; Gaps 5;

Qy 3 QVQLVQSGAEVKKPGASVKASCKGYSNYGIVTWRQAPGGLEWMGWINN--THY 60
Db 1 QVQLQQSGGDLVKPGGSLKVSCAASGFTFSYGMVWRQTPDKRLSEWVAITISGGSYTY 60
Qy 61 AQKQGRVMTTDTSTSTAYMELRSDDTAVYYCAR--DPRGDDP--PYWGQGLTVTVS 117
Db 61 PDSVKGRTISRDNKNTLYLQMSRLKSEDTAMYYCARHINYRQYDGAFDYWGQGLTVTVS 120
Qy 118 SGGGGSGGGGGSEIVLTQSPSLPVTLPGLPQASISCRSSQNLVYS--DGNTYLSWFQQ 176
Db 121 SGGGGSGGGGGGGGGSDIVMAQSPSSLSVAGSKVMSCKSSQSLNSRNRKQYLAQQ 180
Qy 177 RFQGSQPRRLIYKVNRSQVPRDFSGSGSGTDTLTLSRVEAEIDGVYCMQGTWHPPT 236
Db 181 KFGQSPKLIYKASTRESGVPRDFSGSGSGTDTLTLTSSVQAEDLAVYYC--QNDHSYPLT 239
Qy 237 FQGTKEIK 246
Db 240 FGAGTKLEIK 249

RESULT 4
Q65ZC9_HUMAN
ID Q65ZC9_HUMAN PRELIMINARY; PRT; 240 AA.
AC Q65ZC9;
DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Single-chain Fv (Fragment).
GN Name=scFv;
OS Homo sapiens (Human).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP STRAIN=C1G/7;
RC NUCLEOTIDE SEQUENCE.
RX MEDLINE=97362799; PubMed=9219263; DOI=10.1038/nbt0797-629;
RA Kontermann R.E., Wing M.G., Winter G.;
RL Nat. Biotechnol. 15:629-631(1997).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; Y13056; CAA73499.1; -; mRNA.
DR LinkHub; Q652C9; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 2.
KW Immunoglobulin domain.
FT NON_TER 240 240
FT SEQUENCE 240 AA; 25569 MW; FDCFD3645F64B373 CRC64;
SQ
Query Match 52.3%; Score 730.5; DB 2; Length 240;
Best Local Similarity 56.3%; Pred. No. 2.5e-51;
Matches 139; Conservative 47; Mismatches 52; Indels 9; Gaps 4;
OY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNVTGVTWVRQAPQGQLEWGMWGI--NNGNTHY 60
DB 1 QVQLVQSGGVLVQPGGSLRLSCLASGFTFSYGMHWVRQAPGKLEWAVVSYDSNKYI 60
OY 61 AQKPGQRTVMTDTSTSTAYMELSLRSDTAVYYCARDPRGDDPYWGQGTTLTVSSGG 120
DB 61 ADSVKGRTISRDNSKNTLYLQMSLRAEDTAVYYCARD--WCDSLDPWGKGLTVTVSSGG 119
OY 121 GSGGGSGGGGSEIVLTQSPISLTPLVTLGPASISCRSQNLVYSDGNTYLSWFOQRPGQ 180
DB 120 GSGGGSGGGGSDIQMTQSPSTLSASIGDRVTITCRASEGIY-----RWLAWYQQKPKG 174
OY 181 SPRLIYKVNDRSGVDPDRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTWPPRTFGQ 240
DB 175 APKLLIYKASSLASRAPSRFSGSGGTDTFTLTSLQPDPTATYYCQYSNY-PLTFGGG 233
OY 241 TKVEIKR 247
DB 234 TKLEIKR 240
RESULT 5
O7TQM2 MOUSE
ID O7TQM2 MOUSE PRELIMINARY; PRT; 243 AA.
AC O7TQM2
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE ScFv 6H8 protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Balb/c;
RX MEDLINE=22853226; PubMed=12860977; DOI=10.1074/jbc.M306877200;
RA Peter J.C., Eftekhari P., Billiard P., Wallukat G., Hoebeke J.;
```

```
RT "scFv single chain antibody variable fragment as inverse agonist for
RT the beta-2 adrenergic receptor.";
RL J. Biol. Chem. 278:36740-36747(2003).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ574851; CAE00495.1; -; Genomic_DNA.
DR HSSP; P01751; 1A6W.
DR SMR; Q7TQM2; 1-236.
DR MGI; MGI:96514; IGK-V33.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR InterPro; IPR013106; V-set.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 2.
KW Immunoglobulin domain.
FT NON_TER 243 243
FT SEQUENCE 243 AA; 25976 MW; BEFF64D2DCF4F76 CRC64;
SQ
Query Match 51.6%; Score 721.5; DB 2; Length 243;
Best Local Similarity 56.9%; Pred. No. 1.4e-50;
Matches 140; Conservative 43; Mismatches 52; Indels 11; Gaps 4;
OY 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSNVTGVTWVRQAPQGQLEWGMWGI--NNGNTHY 60
DB 1 QVQLVQSGSELVPGASVKLSCKASGYTFTYWMHWVRQKHGQGLEWIGNIYPGSGITNY 60
OY 61 AQKPGQRTVMTDTSTSTAYMELSLRSDTAVYYCARDPRGDDPYWGQGTTLTVSSGG 120
DB 61 DEKFKNKGLITVDTSSSTAYMHLSSLASEDSAVYYCARGGRGLD--VWGAGTTLTVSSGG 118
OY 121 GSGGGSGGGGSEIVLTQSPISLTPLVTLGPASISCRSQNLVYSDGNTYLSWFOQRPGQ 180
DB 119 GSGGGSGGGGSDIQMTQSPSTLSASIGDRVTITCRASEDIY-----NRLAWYQQKPGN 173
OY 181 SPRLIYKVNDRSGVDPDRFSGSGGTDTFTLKISRVEAEDIGVYCMQGTWPPRTFGQ 240
DB 174 APRLLISGNTSLTGVPSRFSGSGGKDYTLTSLQTEDVATYYCQ--YWSRTTGGG 231
OY 241 TKVEIK 246
DB 232 TKLEIK 237
RESULT 6
Q921A6 MOUSE
ID Q921A6 MOUSE PRELIMINARY; PRT; 241 AA.
AC Q921A6
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Anti-CEA 79 single chain Fv (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98170165; PubMed=9509426;
RA Chung J.H., Choi S.J., Kim H.J., Kim I.J., Choi I.H., Lee S.D.,
RA Yi K.S., Suh P.G., Ryu S.H., Chung H.K.;
RT "Cloning and characterization of cDNAs encoding VH and VL of a
RT monoclonal anti-CEA antibody (CEA 79) cross-reactive with NCA-95 and
RT generation of a single-chain Fv molecule (scFv).";
RL Mol. Cells 7:816-819(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=91341421; PubMed=1908510; DOI=10.1084/jem.174.3.613;
RA Stark S.E., Caton A.J.;
```

RT "Antibodies that are specific for a single amino acid interchange in a
RL protein epitope use structurally distinct variable regions.",
J. Exp. Med. 174:613-624(1991).

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

DR EMBL; U88067; AAB48044.1; -; mRNA.

DR PIR; S19965; S19965.

DR PIR; S19967; S19967.

DR PIR; S19968; S19968.

DR PIR; S26325; S26325.

DR HSP; P01607; I8W.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig C2.

DR InterPro; IPR003596; Ig V.

DR InterPro; IPR013106; V-set.

DR SMART; SM00409; IG; 2.

DR SMART; SM00406; IGV; 2.

DR PROSITE; PS00835; IG LIKE; 2.

KW Immunoglobulin domain.

FT NON TER 1 241

FT NON TER 241 241

SQ SEQUENCE 241 AA; 26086 MW; 0276887248B9C771 CRC64;

Query Match 51.4%; Score 718.5; DB 2; Length 241;

Best Local Similarity 54.8%; Pred. No. 2.4e-50;

Matches 138; Conservative 42; Mismatches 53; Indels 19; Gaps 6;

Y 3 QVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGWIN--NGNTHY 60

Y 1 QVQLQSGPELKPGETVKISKASGYTFSTNYGIVTWRQAPGQGLEWMGWINTYGEPTY 60

Y 61 AQKFGQVMTTDTSTAYMELSLRSDDTAVVYCARDPRDDPEYMGQGLTVTVSSGG 120

Y 61 ADPKGAPAFLETSASTAYLQINLNKNEVDATVFCARQDLRLRYFDYMGQGLTVTVSSGG 120

Y 121 GSGSGSGSGSGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVSDGNTYLSWFQORPGQ 180

Y 121 GSGSGSGSGSGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVSDGNTYLSWFQORPGQ 180

Y 121 GSGSGSGSGSGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVSDGNTYLSWFQORPGQ 180

Y 181 SPRL-----IYKVNRRDGVDPFRSGSGSDTFTLKISRVEAEDIGVYCMQGTTHWPP-R 235

Y 176 GPRGAHTLHY----IQGIPSRFSGSGSGRDYSPISNLEPDIATYCL---HYDNLH 228

Y 236 TFGQGTKEIKR 247

Y 229 TFGGTTKLELR 240

RESULT 7

Q65ZL2_9MURI

ID Q65ZL2_9MURI PRELIMINARY; PRT; 487 AA.

AC Q65ZL2;

DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.

DT 11-OCT-2004, sequence version 1.

DT 07-FEB-2006, entry version 13.

DE FV/M4.

GN Name=M4-IPN-<tau>;

OS Mus sp.

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10095;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=96272580; PubMed=8688499;

RA Qi Y., Xiang J.;

RT "A genetically engineered single-gene-encoded anti-TAG72 chimeric

antibody secreted from myeloma cells.";

Hum. Antibodies Hybrids 6:161-166(1995).

CC -----

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

DR EMBL; S82493; AAB37424.2; -; mRNA.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003597; Ig C1.

DR InterPro; IPR003598; Ig C2.

DR InterPro; IPR003006; Ig MHC.

DR InterPro; IPR003596; Ig V.

DR Pfam; PF07654; C1-set; 2.

DR SMART; SM00409; IG; 2.

DR SMART; SM00407; IGV; 1.

DR SMART; SM00406; IGV; 2.

DR PROSITE; PS00835; IG LIKE; 4.

DR PROSITE; PS00290; IG MHC; 1.

SQ SEQUENCE 487 AA; 53578 MW; C7BAB69F30555504 CRC64;

Query Match 51.3%; Score 717.5; DB 2; Length 487;

Best Local Similarity 56.6%; Pred. No. 6.5e-50;

Matches 141; Conservative 46; Mismatches 45; Indels 17; Gaps 7;

Y 2 AQVQLVQSGAEVKKPGASVKVSCKASGYTFSTNYGIVTWRQAPGQGLEWMGWINGN--TH 59

Y 19 SOVQLQQSDAELVKPGASVKISKASGYTFTHAIHWAKQKPGQGLEWIGYISPGNDIK 78

Y 60 YAQKFGQVMTTDTSTAYMELSLRSDDTAVVYCARDPRDDPEYMGQGLTVTVSSG 119

Y 79 YNEKFKGRATLTADKSSSTAYMQLNSLTSEDSAVYFCKSYG---HWGQGTTLT-GSG 133

Y 120 GSGSGSGSGSGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVSDGNTYLSWFQORPG 179

Y 134 GSGSGSGSGSGSEIVLTQSPSLPVLTLGQPASISCRSSQNLVSDGNTYLSWFQORPG 188

Y 180 QSPRLIYKVNRRDGVDPFRSGSGSDTFTLKISRVEAEDIGVYCMQ--GTHWPPRTF 237

Y 189 KSPQLVVAATNLADGVPSRFSGSGSGTQYSLKNSLQSDFGSYVYCOHPWGT---PYTF 245

Y 238 GQGTKEIKR 246

Y 246 GGGTKEIKR 254

RESULT 8

Q925S1_MOUSE

ID Q925S1_MOUSE PRELIMINARY; PRT; 218 AA.

AC Q925S1;

DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.

DT 01-DEC-2001, sequence version 1.

DT 07-FEB-2006, entry version 18.

DE MRP5 (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=BAUB/c;

RX PubMed=11819679;

RA Cui D.X., Zeng G.Y., Wang F., Xu J.R., Ren D.Q., Guo Y.H., Tian F.R.,

Yan X.J., Hou Y., Su C.Z.;

"Mechanism of exogenous nucleic acids and their precursors improving

the repair of intestinal epithelium after gamma-irradiation in mice.";

World J. Gastroenterol. 6:709-717(2000).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=BAUB/c;

RA Cui D., Zeng G., Yan X., Li X., Su C.;

"Cloning of mouse genes related to repairing of intestinal epithelium

of the irradiated mice by treatment with the intestinal RNA of mice of

the same strain.";

RT the same strain.";

CC -----

Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80 (2001).
 Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AF240168; AAK43733.1; -, mRNA.
 HSSP; P01665; 1QNZ.
 InterPro; IPR003599; IG.
 InterPro; IPR007110; IG-like.
 InterPro; IPR003596; IG.v.
 InterPro; IPR013106; V-set.
 SMART; SM00409; IG; 1.
 SMART; SM00406; IGv; 1.
 PROSITE; PS50835; IG LIKE; 1.
 Immunoglobulin domain.
 NON_TER 218
 SEQUENCE 218 AA; 23013 MW; 527E4FA8F7982817 CRC64;

Query Match 47.5%; Score 664.5; DB 2; Length 218;
 Best Local Similarity 57.5%; Pred. No. 5.3e-46;
 Matches 126; Conservative 42; Mismatches 48; Indels 3; Gaps 2;

1 MAQVQLVQSGAEVKKPGASVKVSCKASGYTFSNYGITWVRQAPGQLEWGMGIN--NGNT 58
 1 MAQVQLQSGPELKPKGTWRISCKASGYTFTAGQWQKMPGKGLKWIWINTHSGVP 60

59 HYAQKFGQRTMTTDTSTAYMELRLSRDDTAVYVCARDPRGDEPYWGQGLTVTVSS 118
 61 KYAEFGKGRFAFSLTASTAYLQISNLKEDTATYFCWRWDYDGGPAYWGQGLTVTVSS 120

119 GGGSGGGGGGGGSEIVLTQSPSLPVTLPQASISCRSSQNLVYSDGNTYLSWFOQRP 178
 121 GGGSGGGGGGGSDIVLTQSPASLAVSLGQRATISCRASES-VDNIGISFMNWFQKRP 179

179 GQSPRLIYKVNDRDGVDFRFGSGSGTDFTLKISRVE 217
 180 GQPKLLIYAASQGGVPGAGLIASGSGTDFSLNIYPME 218

RESULT 9
 Q9UL80 HUMAN
 ID Q9UL80 HUMAN PRELIMINARY; PRT; 114 AA.
 AC Q9UL80;
 DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
 DT 01-MAY-2000, sequence version 1.
 DT 07-FEB-2006, entry version 21.
 DE Myosin-reactive immunoglobulin light chain variable region (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98271139; PubMed=9614934; DOI=10.1006/clin.1998.4531;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus.";
 RL Clin. Immunol. Immunopathol. 87:184-192 (1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92352481; PubMed=1322670;
 RA Stuber F., Lee S.K., Bridges S.L. Jr, Koopman W.J., Schroeder H.W. Jr,
 RA Gaskin F., Fu S.M.;
 RT "A rheumatoid factor from a normal individual encoded by VH2 and V kappa II gene segments.";
 RL Arthritis Rheum. 35:900-904 (1992).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=93170387; PubMed=8436174;
 RA Wagner S.D., Luzzatto L.;
 RT "V kappa gene segments rearranged in chronic lymphocytic leukemia are

distributed over a large portion of the V kappa locus and do not show somatic mutation";
 Eur. J. Immunol. 23:391-397 (1993).
 [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92289816; PubMed=1601042;
 RA Huber C., Klobeck H.G., Zachau H.G.;
 RT "Ongoing V kappa-J kappa recombination after formation of a productive V kappa-J kappa coding joint.";
 Eur. J. Immunol. 22:1561-1565 (1992).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AF035034; AAD56270.1; -, mRNA.
 PIR; B49002; B49002.
 PIR; S23638; S23638.
 PIR; S34094; S34094.
 PIR; S34095; S34095.
 HSSP; P01625; LLVE.
 SMR; Q9UL80; 1-114.
 InterPro; IPR003599; IG.
 InterPro; IPR007110; IG-like.
 InterPro; IPR003596; IG.v.
 InterPro; IPR013106; V-set.
 SMART; SM00409; IG; 1.
 SMART; SM00406; IGv; 1.
 PROSITE; PS50835; IG LIKE; 1.
 Immunoglobulin domain.
 NON_TER 114
 NON_TER 114
 SEQUENCE 114 AA; 12775 MW; 070E31E210D1CB01 CRC64;

Query Match 40.4%; Score 565; DB 2; Length 114;
 Best Local Similarity 92.1%; Pred. No. 3.2e-38;
 Matches 105; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

134 EIVLTQSPPLSLPVTLPQASISCRSSQNLVYSDGNTYLSWFOQRPQSPRLIYKVSNRD 193
 1 DVVMTQSPPLSLPVTLPQASISCRSSQSPVSDGNTYLSWFOQRPQSPRLIYKVSNRD 60

194 SGVPRFRFGSGSGTDFTLKISRVEADIGVYCMQGTHTWPPRTFGQTKVEIKR 247
 61 SGVPRFRFGSGSGTDFTLKISRVEADVGIVYCMQGTHTWPPRTFGQTKVEIKR 114

RESULT 10
 KV2F HUMAN
 ID KV2F HUMAN STANDARD; PRT; 133 AA.
 AC P06310;
 DT 01-JAN-1988, integrated into UniProtKB/Swiss-Prot.
 DT 01-JAN-1988, sequence version 1.
 DT 07-MAR-2006, entry version 41.
 DE IG kappa chain V-II region RPMI 6410 precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RX MEDLINE=86041852; PubMed=2997711;
 RA Klobeck H.G., Meindl A., Combratio G., Solomon A., Zachau H.G.;
 RT "Human immunoglobulin kappa light chain genes of subgroups II and III.";
 RL Nucleic Acids Res. 13:6499-6513 (1985).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; Z00020; CAA77315.1; -, Genomic DNA.
 PIR; A01890; K2HURP.
 HSSP; Q99M37; 1191.

```
DR SMR; P06310; 21-133.
DR Ensembl; ENSG00000173758; Homo sapiens.
DR LinkHub; P06310; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07686; V-set; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 133
FT REGION 21 43
FT REGION 44 59
FT REGION 60 74
FT REGION 75 81
FT REGION 82 113
FT REGION 114 122
FT REGION 123 132
FT DISULFID 43 113
FT NON TER 133 133
SQ SEQUENCE 133 AA; 14707 MW; 513CCAF3673009EE CRC64;

Query Match 40.1%; Score 560.5; DB 1; Length 133;
Best Local Similarity 89.8%; Pred.No. 8.8e-38;
Matches 106; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Qy 130 GGGSEIVLTQSPVLTGQPASISCRSSQNLVYSGNTYLSWFGQRPQGSPRLIYKV 189
Db 17 GSSGDVMTQSPVLTGQPASISCRSSQSLVYSGNTYLNWFPQRPQGSPRLIYKV 76
Qy 190 SNRDGVPDRFSGSGGTDFTLTKISRVAEIDGVYCMQGHWPRTFGQGTQVEIKR 247
Db 77 SNRDGVPDRFSGSGGTDFTLTKISRVAEIDGVYCMQGHWSWTFGQGTQVEIKR 133

RESULT 11
Q8TCD0 HUMAN
ID Q8TCD0 HUMAN PRELIMINARY; PRT; 239 AA.
AC Q8TCD0.
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J.J., Heiton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
```

```
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92285150; PubMed=1598223;
RA Hirabayashi Y., Munakata Y., Sasaki T., Sano H.;
RT "Variable regions of a human anti-DNA antibody O-81 possessing lupus
nephritis-associated idiotype.";
RL Nucleic Acids Res. 20:2601-0(1992).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92201291; PubMed=1551402;
RA Lautner-Rieske A., Huber C., Meindl A., Pargent W., Schable K.F.,
Thiele R., Zocher I., Zachau H.G.;
RT "The human immunoglobulin kappa locus. Characterization of the
duplicated A regions.";
RL Eur. J. Immunol. 22:1023-1029(1992).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94080891; PubMed=8258341;
RA Klein R., Jaenichen R., Zachau H.G.;
RT "Expressed human immunoglobulin kappa genes and their hypermutation.";
RL Eur. J. Immunol. 23:3248-3262(1993).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93170387; PubMed=8436174;
RA Wagner S.D., Luzzatto L.;
RT "v kappa gene segments rearranged in chronic lymphocytic leukemia are
distributed over a large portion of the v kappa locus and do not show
somatic mutation.";
RL Eur. J. Immunol. 23:391-397(1993);
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC022362; AAH22362.1; -; mRNA.
DR PIR; S22658; S22658.
DR PIR; S34095; S34095.
DR PIR; S40324; S40324.
DR PIR; S40374; S40374.
DR PIR; S42267; S42267.
DR PIR; S42268; S42268.
DR HSP; P01834; I172.
DR SMR; Q8TCD0; 21-237.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR013106; V-set.
DR Pfam; PF07654; C1-set; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00407; IGC1; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26235 MW; FACEDC3A3B0387ID CRC64;

Query Match 39.9%; Score 557.5; DB 2; Length 239;
Best Local Similarity 87.3%; Pred.No. 3.1e-37;
Matches 103; Conservative 10; Mismatches 4; Indels 1; Gaps 1;

Qy 130 GGGSEIVLTQSPVLTGQPASISCRSSQNLVYSGNTYLSWFGQRPQGSPRLIYKV 189
```

	Best Local Similarity	78.8%;	Pred. No.	5.1e-32;	Matches	93;	Conservative	14;	Mismatches	10;	Indels	1;	Gaps	1;
Qy	130	GGGSEIVLTOSPLSLPVTLLQPASISCRSSQNLYVSDNTYLSWFOQPQCSPRRILIYKV	189	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Dd	17	GSSGDIVMTQTPLSSPVTLGPASISCRSSSELHNSGNTYLSWLHQRPQPRLLIYKI	76	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Qy	190	SNRDSGVDPFRFSGSGGTDTFLKISRVEADICVYYCMQGTHWPPTFFGGTKVEIKR	247	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Dd	77	SNRFSGVDPFRFSGSGAGTDTFLKISRVEADVGYVCMVSHF-PRTFGQGRVEIKR	133	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
RESULT 13														
ID	Q6N091	HUMAN	PRELIMINARY; PRT; 500 AA.											
AC	Q6N091;													
DT	05-JUL-2004,	integrated into UniProtKB/TrEMBL.												
DT	05-JUL-2004,	sequence version 1.												
DT	07-FEB-2006,	entry version 12.												
DE	Hypothetical protein DKFZp686C0220 (Fragment).													
GN	Name=DKFZp686C0220;													
OS	Non sapiens (Human)													
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;													
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;													
OC	Homio.													
OX	NCBI_TaxID=9606;													
RN	[1]													
RC	NUCLEOTIDE SEQUENCE.													
RG	TISSUE=Human rectum tumor;													
RA	The German Human cDNA Consortium;													
RA	Wambutt R., Heubner D., Newes H.W., Weil B., Amid C., Osanger A.,													
RA	Fobo G., Han M., Wiemann S.;													
RC	Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.													
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms													
CC	Distributed under the Creative Commons Attribution-NoDerivs license													
CC	-----													
DR	EMBL; BX640625; CAE45779.1; -; mRNA.													
DR	HSBP; P01751; IA6W.													
DR	SMR; Q6N091; 270-478.													
DR	InterPro; IPR003599; Ig.													
DR	InterPro; IPR007110; Ig-like.													
DR	InterPro; IPR003597; Ig.cl.													
DR	InterPro; IPR003006; Ig_MHC.													
DR	InterPro; IPR003596; Ig_v.													
DR	InterPro; IPR013106; V-set.													
DR	Pfam; PF07654; Cl-set; 2.													
DR	SMART; SM00409; IG; 1.													
DR	SMART; SM00407; IGc1; 2.													
DR	SMART; SM00406; IGv; 1.													
DR	PROSITE; PS50835; IG_LIKE; 4.													
DR	PROSITE; PS00230; IG_MHC; UNKNOWN 2.													
KW	Hypothetical protein; Immunoglobulin domain; Repeat.													
FT	NON TER													
SQ	SEQUENCE	500 AA;	54160 MW;	3C423A17D65A41E4	CRC64;									
Query Match														
35.3%; Score 493.5; DB 2; Length 500;														
Best Local Similarity 56.1%; Pred. No.1.2e-31;														
Matches 106; Conservative 15; Mismatches 27; Indels 41; Gaps 5;														
Qy	2	AQQLVQSQAEEKKPGASVKVSCKASYGTFPSNYGITWTWRQAPQGLEWMGWIN--NGNTH	59	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Dd	37	SQLVLVQSQAEEKKPGASVKVSCKASYGTFSDSHITLWRQAPQGLEWIGWSAYSGQTY	96	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Qy	60	YAQKFGRVTMTDTSTSTAYNELSLRSDDTAVVYCARDPR-----GDDEFYWGQGITLV	114	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Dd	97	YAQNLQGRVTMTDTSTSTAYNELSLRSDDTAVVYCARKDQSYTTIPNDAPHIWGGQTMV	156	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Qy	115	TVSSSGGGSGGGSGGGSEIVLTOSPILPVTLCGPASISCRSSQNLYVSDNTYLS--	172	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Dd	157	TVSSAS-----PTSPKVPFPULDSTP-----QDNVVVAACL	187	:	: :	: :	: :	: :	: :	: :	: :	: :	: :	: :
Qy	173	--WFQRP	178											

[illegible]

Search completed: August 9, 2006, 13:14:39
Job time : 303 secs